ACOUSTIC AND KINEMATIC CORRELATES OF PHONOLOGICAL LENGTH CONTRAST IN ITALIAN CONSONANTS

Barbara Gili Fivela, Claudio Zmarich, Pascal Perrier^, Christophe Savariaux°, Graziano Tisato**

Università del Salento Lecce-Italy, *ISTC-CNR Padova-Italy, ^ICP-INPG Grenoble-France, °ICP-Université Stendhal, Grenoble-France
barbara.gili@ateneo.unile.it, zmarich@pd.istc.cnr.it, perrier@icp.inpg.fr, christophe.savariaux@icp.inpg.fr, gtisato@pd.istc.cnr.it

ABSTRACT

In Italian, length contrast is exploited in the consonant system. Previous articulatory studies have focused on the temporal organization of gestures in Italian geminates and on the lower lip kinematics of the singleton/geminate distinction. In this paper, data on lip and tongue gestures are discussed in order to directly test hypotheses on the gestural organization of geminate consonants and to collect observations on the possible position of gestural targets in geminate and singleton consonants.

The results show that Italian geminates appear to be best accounted for by a hybrid model that makes use of both Öhman's Vowel-to-Vowel model and Browman and Goldstein's Vowel-and-Consonant model. Moreover, our data partly confirm the existence of a higher virtual target in geminates than in singletons.

Keywords: length contrast, geminates, kinematics, Italian, consonants.

1. INTRODUCTION

The phonetic characteristics of geminate consonants in Italian have been reported in a number of papers, especially as far as acoustic and auditory features are concerned. Various studies show that geminate consonant duration is about twice singleton duration, and that the vowel immediately preceding a geminate consonant is shorter than the vowel preceding a singleton [1, 3]. The first kinematic study taking account of Italian geminates [8] focused on the temporal organization of consonantal gestures with respect to vowels and showed that Italian geminates are better represented by the Vowel-to-Vowel model proposed by Öhman [7] than by the Vowel and Consonants model of Browman and Goldstein [2]. Other studies [10, 4], focusing on the monosegmental vs. bisegmental representation of geminates and on the kinematics of lower lip gesture, have provided evidence for the heterosyllabic representation of geminate consonants, due to their similarity with heterosyllabic clusters. In fact, kinematic measurements show that the opening gesture has longer duration, greater amplitude and time-to-peak, and lower stiffness in both geminates and cluster segments than in singleton consonants.

This work stems from the above-mentioned comparisons of Italian geminate and singleton consonants. The goal of this paper is to discuss data on a wider range of gesture types (both lip and tongue gestures), and to directly test some hypotheses concerning the gestural organization of geminate consonants. First of all, the aim is to directly test Smith's hypothesis [8] on the temporal organization of gestures in Italian geminate consonants, and secondly, to collect observations on the possible position of gestural targets in geminates and singleton consonants. We will consider Löfqvist's [5] suggestion that a higher virtual target could underlie the production of geminate as compared to singleton consonants.

2. GESTURE TEMPORAL ORGANIZATION AND GESTURE TARGET

According to Smith [8], in syllable-timed languages such as Italian, the time interval between the nuclei of two successive syllables does not depend on the number of intervening consonants. Smith's data from three Italian speakers, producing bilabial consonants in an /i/-to-/a/ vowel context, support the hypothesis that in geminate consonants the closure gesture is anticipated during the preceding vowel and the opening gesture is postponed during the following vowel, while the centre of maximum constriction is timed similarly for both types of consonant. According to Smith, the model which best accounts for vowel and consonant gesture timing in Italian is Öhman's [7] Vowel-to-Vowel model, where the consonant gesture is considered to be superimposed on a vowel-to-vowel gesture which determines the basic timing. In contrast, in a language such as Japanese, which is not syllable-timed, temporal gesture organization seems to be best represented by

Browman and Goldstein's [2] Vowel-and-Consonant model. Indeed, the vowel gesture starts at the end of the consonantal one, meaning that both the time interval between the two vocalic nuclei and the timing of maximum consonantal constriction depend on the number of intervening consonants (for Japanese, see also [6]). Smith's hypothesis was tested for Italian by Zmarich and Gili Fivela [10] by considering bilabial nasals realized in an /a/-context. One out of two subjects appeared to adhere to the Vowel-to-Vowel model, while for the other, Browman and Goldstein's model better suited the data.

As for the geminate/singleton gesture target, in a study of bilabial plosives and fricatives Löfqvist [5] proposed that differences in lip contact duration could originate both from the higher position of the virtual target of the geminate and from its delayed position. Given the correlation between movement displacement and velocity peak, the author tested the hypothesis that speakers could control the closure-constriction duration by varying the position of the virtual consonant target. Specifically, he looked at the velocity at the closure, which is expected to be greater for higher virtual targets. However, the data on the lower lip kinematics of Swedish and Japanese speakers producing singleton and geminate consonants showed that there is no significant difference in velocity peak in long and short consonants (apart from one Swedish subject). Hence, Löfqvist [5] concluded that subjects do not only change the virtual target position, but also vary its timing, i.e. they change the deceleration rate in order to keep the contact longer.

In this work, we will test Smith's and Löfqvist's hypotheses in both normal and fast speech production in Italian, as a higher speech rate often reduces the phonetic contrast of two structures in phonological opposition, and reduces the number of alternative articulatory strategies.

3. CORPUS AND METHOD

The corpus exploited for the investigation was composed of words and pseudo-words containing bilabial, dental and velar consonants realized as singletons (/m, l, d, g/), geminates (/mm, ll, dd, gg/) and cluster segments (/ld, dl/) in an /a/-context (e.g., 'mama', 'mamma', 'lala', 'lalla', 'lalda', 'ladla'); velars are not considered here. For bilabial nasals only, an /i/-to-/a/ context was also considered (e.g., 'mima', 'mimma'), in order to check Smith's hypothesis on the basis of completely comparable data.

Target words were inserted in a carrier sentence, before an adverb whose first consonant was the same as the consonant in the target word (eg. 'malamente' for *mama*, *mamma*, etc.: , *chiama mama malamente* "s/he calls mama badly"). Speakers were asked to read each sentence ten times both at a natural speech rate and, immediately afterwards, at a faster speech rate. Four speakers were recorded, but only two of them are considered in this paper: a female speaker of a north-eastern variety of Italian (NE) and a female speaker of a north-western variety (NW). The audio signal was acquired by means of a DAT recorder, while the kinematic data were collected using the 2D EMA system at the ICP lab in Grenoble. As well as the sensors used for reference (glued to the nose and the upper incisors), two were placed on the upper and lower lips, and four were glued to the midsagittal plane of the tongue in the range of about 1 cm to 5 cm from the tongue apex. The data discussed here only relate to measurements by the lower lip and tongue apex sensors for consonants, and by the tongue dorsum sensor for vowels.

3.1. Auditory test

Consonant duration is distinctive in (standard) Italian, but various dialects spoken in northern Italy do not use geminates and do not exploit the geminate/singleton contrast. First of all then, we decided to verify, by means of a perception test, whether our speakers had adequately produced geminate consonants. Secondly, we sought to look at acoustic/articulatory correlates of clear geminate *vs.* singleton contrasts.

A selection of the 480 stimuli produced by the two subjects was used for the perception test (for each subjects, two repetitions realized at a normal speech rate and all repetitions realized at fast rate were selected - 240 in total). Five subjects from the northern part of Italy (Torino) and five subjects from the southern one (Lecce-Taranto) took part in the perception test. Using *Perceval* software (developed at LPL, Aix-en-Provence), subjects listened to audio files containing target words and judged whether the words included singleton or geminate consonants. At this stage, the test results were used in order to select unambiguous cases of singleton/geminate contrast. Cases in which the stimuli were wrongly recognized by at least 4 subjects during the perception test were not considered for acoustic and kinematic analysis.

The comparison of these ‘no contrast’ cases (mainly observable at a faster speaking rate) with the ‘clear contrasts’ was left to a follow-up study.

3.2. Measurements

The duration of the consonant and both the preceding and the following vowels was measured by manually segmenting and labelling the acoustic signal. Kinematic measurements were performed after semiautomatic segmentation and labelling of the signal [9]. The duration, amplitude and velocity of lower lip/tongue apex gestures (and tongue dorsum gestures for bilabials in an /i/-to-/a/ context) were calculated. For the closing gesture, for instance, the duration was calculated as the time interval between the maximum opening and the maximum closure of the articulators; the amplitude and velocity were calculated, respectively, as the vertical component of the articulator displacement and as the maximal peak on the velocity curve during the gesture.

4. GESTURE TEMPORAL ORGANIZATION

One-way ANOVAs were performed on the normal and fast speech rate recordings, checking whether the geminate/singleton factor was significant in relation to the vowel-consonant phasing.

The results show that there are no significant differences in the time interval from the highest tongue dorsum position for [i] to the bilabial closure for [m] and [mm], and this was seen in both subjects, in both normal-rate production (NE: [F(1,17)=1.992; p>0.05], NW: [F(1,17)=0.345; p>0.05]), and in fast-rate production (NE: [F(1, 17)=0.003; p>0.05], NW: [F(1, 14)=0.205; p>0.05]). However, the time interval from the bilabial closure for [m] and [mm] to the lowest tongue dorsum position for [a] is greater for both subjects’ geminates in normal-rate production, although significance is seen only with the NW subject ([F(1,17)=0.492; p>0.05], [F(1,17)=109.59; p<0.01]); at the fast rate, the difference is significant only for the NE subject ([F(1,17)=6.867; p<0.05]; NW: [F(1,14)=0.002; p>0.05]). Moreover, the vowel-to-vowel time interval, i.e. from the highest tongue dorsum position for [i] to the lowest tongue dorsum position for [a], is significantly greater in geminates for both subjects in normal-rate production (NE: [F(1,17)=4.435; p<0.05], NW: [F(1,17)=49.027; p<0.05]) and for the NE subject in fast-rate production ([F(1,17)=8.859; p<0.05], NW: [F(1,14)=0.000; p>0.05]). In fact, at the normal speech rate both the closing and the opening labial gestures are longer in geminates than in singletons, for both subjects (closing and opening for NE: [F(1,18)=19.619; p<0.01], [F(1,18)=32.794; p<0.01] respectively; for NW: [F(1,17)=12.405; p<0.01], [F(1,17)=140.753; p<0.01]). At the fast rate, only the closing gesture is significantly different, and only for the NE subject ([F(1,17)=34.297; p<0.01]). Thus, Smith’s hypothesis of the constant timing of maximum constriction with respect to tongue movement is not completely confirmed by our data, at least at the normal speech rate.

However, the interval between the highest tongue dorsum position for [i] to the starting point of bilabial closure is significantly shorter in geminates than in singletons for both subjects at the normal speech rate (NE: [F(1,17)=4.450; p<0.05], NW: [F(1, 17)= 15.994; p<0.05]), and for neither subject at the faster rate (NE: [F(1, 17)=3.656; p>0.05], NW: [F(1, 14)=0.058; p>0.05]). In contrast, the interval between the end point of bilabial closure and the lowest tongue dorsum position for [a] is not significantly different in geminate and singleton consonants, for either subject or speech rate (normal and fast for NE: [F(1,17)=0.998; p>0.05], [F(1,17)=3.791; p>0.05] respectively; for NW: [F(1,17)=1.151; p>0.05], [F(1,14)=0.008; p>0.05]). Consistent results are obtained by means of analyses inspired by the recent work of Löfqvist [6]. Similarly to what happens in case of Japanese speakers, for our subjects the onset and offset of the vowel-to-vowel tongue movement occur before and after the acoustic bilabial closure respectively, for both geminates and singletons. However, only for our subjects does the beginning of the closure occur significantly earlier for geminates than for singletons (normal and fast rate for NE: [F(1,17)=17.583; p<0.05], [F(1,17)=9.804; p<0.05] respectively; NW: [F(1,16)=11.069; p<0.05], [F(1,14)=0.605; p>0.05]), while the interval from the end of bilabial closure to the end of the tongue movement does not reveal any significant difference. Thus Smith’s hypothesis concerning the anticipation of the geminate consonantal gesture in the preceding vowel and the delay in the following vowel is confirmed only for the former.

5. GESTURE VIRTUAL TARGET

In order to test Löfqvist's [5] hypothesis, the velocity peak position was checked. One-way ANOVAs were performed in order to check for the influence of the geminate/singleton factor on the magnitude of the velocity curve at the onset and offset of the acoustic closure, and on the temporal position of the velocity peak with respect to the acoustic boundaries.

The results for both magnitude and position are summarized in Table 1 (upper table for the normal rate results, and lower table for the fast rate), where 'Y(es)' means significant ($p<0.05$). For the opening gesture in alveolar consonants, the velocity curve at closure release has a significantly higher value in geminates, and its maximum peak is temporally closer to the release in geminates than in singletons (as seen for bilabials with a Swedish speaker referred to in [5]); at the fast rate, this is true for the NE subject in all segments. In contrast, for the closing gesture at the normal rate, only few significant results are found for the two measures, and never for the same subject; at the fast rate, the results are always significant only for the NE subject.

Table 1: Summary of one-way ANOVAs results: Normal rate (upper table) and fast rate (lower table) for both subjects. Y(es) indicates a significant ($p<0.05$) difference between singletons and geminates for the magnitude of the velocity curve at the onset and offset of closure (left) and for the temporal lag of its max peak with respect to the onset and offset of closure (*right*) in bilabial nasals and alveolar laterals and plosives.

	Blb.Nas				Alv.Lat and Plos.			
	NE		NW		NE		NW	
Closing	N	*Y*	Y	*Y**	N	*N*	N	*N*
Opening	Y	*N*	N	*N*	Y	*Y*	Y	*Y*

	Blb.Nas				Alv.Lat and Plos.			
	NE		NW		NE		NW	
Closing	Y	*Y*	N	*Y*	Y	*Y*	N	*N*
Opening	Y	*Y*	Y	*N*	Y	*Y*	N	*N*

*shorter lag for singletons than for geminates

6. DISCUSSION AND CONCLUSIONS

Smith's results [8] appear not to be completely confirmed at the normal speech rate. Our data show that the timing of maximum constriction with respect to tongue movement was not always constant in geminate\singleton consonants, and the vowel-to-vowel interval varied significantly (this tends to fit better with Browman and Goldstein's model rather than with Öhman's). On the other hand, Smith's hypothesis concerning the anticipation of the geminate consonantal gesture in the preceding vowel is confirmed (consistent with Öhman's model); however, no significant delay in the following vowel is observed. At the fast speech rate, geminates and singletons show few significant differences, limited to a single subject (NE). Thus Italian geminates appear to be best accounted for by a hybrid of Öhman's Vowel-to-Vowel model and Browman and Goldstein's Vowel-and-Consonant model, that is, it needs to account for the following: the variable duration of the vowel-to-vowel gesture, the only partially constant timing of maximum consonantal constriction, and the anticipation of the geminate consonantal gesture in the preceding vowel.

As for the geminate/singleton gesture target, Löfqvist's [5] initial hypothesis of a higher virtual target for geminates than for singletons is partly confirmed for the articulators considered here: At the normal rate, the hypothesis is verified in the opening gesture of alveolar consonants; at the fast rate, it verified for all types of consonants in one subject.

7. REFERENCES

[1] Bertinetto, P.M. 1981. *Strutture prosodiche dell'italiano*. Accademia della Crusca, Firenze.

[2] Browman C.P. & Goldstein L.M. 1986. Towards an articulatory phonology. *Phonology Yearbook, 3*, 219-252.

[3] Esposito, A. & Di Benedetto M.G. 1999, Acoustical and perceptual study of gemination in italian stops, JASA, 106, 4, 2051-2062.

[4] Gili Fivela, B., Zmarich, C. 2005. Italian Geminates under Speech Rate and Focalization Changes: Kinematics, Acoustic, and Perception Data. *InterSpeech 2005*, Lisbon, 2897-2900.

[5] Löfqvist, A. 2005. Lip kinematics in long and short stop and fricative consonants. *JASA*, 117, 2, 858-878.

[6] Löfqvist, A. 2006. Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese. *JASA*, 120, 5, 2872-2883.

[7] Öhman S.E.G. 1967. Numerical Model of Coarticulation. *JASA., 41*, 310-320.

[8] Smith, C.L. 1995. Prosodic patterns in the coordination of vowel and consonant gestures. In: B.Connell & A.Arvaniti (eds) *Papers in Laboratory Phonology IV, Phonology and phonetic evidence*. CUP, 205-222.

[9] Tisato G., Cosi P., Sommavilla G., Zmarich C. In press, Recent improvements in *InterFace* software: New tools for Articulatory Signal Processing. *Atti del III° Convegno Nazionale AISV*, Trento.

[10] Zmarich, C., Gili Fivela, B. 2005. Consonanti scempie e geminate in italiano: studio cinematico e percettivo dell'articolazione bilabiale e lebiodentale. *Atti del I° Convegno Nazionale AISV*, Padova, 429-448.