ACOUSTIC ANALYSIS OF ITALIAN SINGLETON/GEMINATE STOP PRODUCTION IN TWO AMBIENT TEMPERATURE CONDITIONS

Lia Saki Bučar Shigemori, Alessandro Vietti

ALPS - Alpine Laboratory of Phonetic Sciences, Free University of Bozen - Bolzano
LiaSaki.BucarShigemori@unibz.it

ABSTRACT

Italian is subject to regional variation in the phonetic realization of phonological features [2]. This study examines the duration of the phonemic quantity and voicing contrast of Italian stop consonants produced by speakers from South Tyrol. Speakers were recorded once in a soundproof booth and once in a cold chamber to additionally test whether speech patterns were affected by cold temperature. This allows us to test the stability of patterns observed and to investigate whether temperature has a systematic effect on speech as is the case for speech under other kinds of stress. The main correlates of the singleton/geminate contrast were stop closure and preceding vowel duration, in line with previous studies on Standard Italian [4]. Release duration was longer for velar consonants compared to apical and bilabial stops, consistent with results reported for voiceless geminates [18]. No systematic temperature effect was observed.

Keywords: geminates, Italian, acoustic duration, temperature, environmental effects

1. INTRODUCTION

The aim of this study is to investigate whether phonological contrasts remain stable at cold temperatures. Stress factors such as loud background noise, emotion or fatigue affect different parameters of speech production, including pitch, duration, voice quality and formant values [6, 16]. Cold ambient temperature might be expected to induce stress not only on the emotional but also on the physical level. It is known to lead to muscle rigidity [20], which might affect lip and jaw movement during speech production. Low temperature is generally believed to affect vocal fold function which might lead to vocal injury, for example of singers performing outdoors. However, empirical studies are limited to measurements of phonation threshold pressure and perceived phonatory effort [14]. Although the immediate effect of ambient temperature on speech production has to our knowledge not been studied, several typological studies correlating the ecological environment and phonetic inventories have been published [11, 5]. In this study we examine the stop production in Italian speakers from South Tyrol, a province in northern Italy. Italian is known as a true voicing language, in which vocal fold vibration is sustained during voiced stops in contrast to languages where the phonological voicing contrast is phonetically realized by difference in voice onset time after the burst. Furthermore, stops in Italian can be realized as singletons or geminates; thus duration is phonological. To maintain voicing, a pressure drop across the vocal folds is required [23]. Therefore, speakers might expand the size of the cavity between the oral closure and the larynx by lowering the larynx, fronting the place of the closure, lowering the jaw or relaxing the cheek [9]. Assuming that phonation in general becomes more difficult at cold temperatures, we hypothesize that voicing for geminate stops is also more difficult to maintain. In addition, bilabials might pattern differently than velars or alveolars, because the articulators involved in their production are more exposed to ambient temperature. In this paper we limit the analysis to acoustic duration of stop closure, release duration and preceding vowel, focussing on whether voiced segments and bilabials are particularly affected by cold ambient temperature. The main correlate of the Italian singleton/geminate contrast is consonant duration. While consonant duration differs for different stop categories, the quantity contrast is also affected by different prosodic factors [4, 12]. Additionally, the preceding vowel duration has been found to be shorter when the consonant is a geminate compared to when it is a singleton [4]. Descriptive literature on Italian claims that stop consonants in Italian are generally unaspirated. While some studies measured the consonant duration as including also aspiration [12], in those studies that measured aspiration separately it was affected by consonant category, vowel context and speakers' origin [4, 18]. According to descriptive sources, the quantity contrast is less salient in northern varieties due to degemination of long consonants [2, 22]. While South Tyrol lies in northern Italy, the variety of Standard Italian spoken in the region is said to be less regionally marked and more standard-like [21]. However, acoustic and articulatory studies of Italian spoken in South Tyrol are scarce and to our knowledge the durational properties of stop consonants and the singleton/geminate contrast have not been studied yet.

The main aim of this study is to assess whether durational contrasts are maintained in different temperature conditions, but it will also provide new data on geminate production by Italian speakers from South Tyrol.

2. METHODS

Speech of three female Italian speakers from South Tyrol was analyzed. Each subject was recorded first in a soundproof booth (warm condition, 24°C - 29°C) and on a separate day in a cold chamber (cold condition, 2°C - 4°C). Simultaneous recordings using electroglottography and ultrasound were carried out, but this data has not been examined yet. Participants read sentences as they were presented on a screen one at a time. The target consonants were singleton and geminate bilabial, alveolar and velar stops contrasting in voicing. Target words were real words in which the target consonant appeared in one of two vocalic contexts: always preceded by lexically stressed /a/, but in one case followed by /a/ and in the other followed by /i/. Target words were presented in two sets of more or less meaningful carrier sentences, but in this paper we will analyze only those in which the target word was the left-dislocated object as in:

I **dati** non li ho raccolti io.

(The data, I have not collected it.)

Within each session each sentence was repeated four times. Reading mistakes and tokens with loud background noise were removed, resulting in 549 tokens (see Tab. 1). The acoustic data were automatically segmented using WebMAUS [7], manually corrected in Praat [3] and converted to an EmuR database [24]. Within the stop consonants closure and release (burst and aspiration) were annotated as separate segments. The release duration differs from VOT, in that for voiced and voiceless stops it was measured from the end of complete closure to the beginning of periodicity of the following vowel, regardless of whether voicing was present during closure or not.

For consonant closure duration, preceding vowel duration and release duration three separate linear mixed effect models were fitted using the lmerTest-package [8] in R [13]. The models consisted of phonemic QUANTITY (singleton, geminate), VOICING (voiced, voiceless), PLACE of articulation (bilabial, alveolar, velar) and TEMPERATURE (warm, cold) as fixed factors and SPEAKER (with by-SPEAKER intercept and slope for QUANTITY, VOICING, PLACE and TEMPERATURE), REPETITION and ITEM as random factors. The significance of each effect was retrieved from the ANOVA type III table and *p*-values were calculated using Satterthwaite's method of approximation. We considered an effect to be significant when $p < 0.05$. Estimate means calculations and post-hoc Tukey-tests were carried out using the emmeans-package [10].

Table 1: Numbers of tokens analyzed.

	voiceless singleton		geminate		voiced singleton		geminate	
	cold	warm	cold	warm	cold	warm	cold	warm
alveolar	21	23	22	24	21	24	21	24
bilabial	23	24	23	24	24	24	24	24
velar	22	24	23	24	23	24	19	20

3. RESULTS

3.1. Closure duration

The closure durations based on actual measurements obtained from all three speakers are illustrated in Fig. 1. Upon visual inspection, the duration differs by place of articulation, and the singleton/geminate contrast seems robust with geminates being about twice as long as singletons. In most cases the closure is longer when the stop is voiceless than voiced. There seem to be no systematic differences between the cold and warm conditions. The statistics confirm what can be observed in Fig. 1: Main effects of QUANTITY (F[1, 3.18]= 34.5, p<0.01), VOICING (F[1,7.25]=32.1, p<0.01) and PLACE (F[2,6.01]=9.5, p=0.05) were significant. Our main interest was the effect of TEMPERATURE, which was not significant as a main effect (F[1,3.0]=2.3, p=0.22), but several interaction effects, in particular together with place of articulation were significant. The estimate means of the closure durations for each phoneme in the two temperature condition as predicted by the full model are presented in Tab. 2. Because we were not primarily interested in durational differences between places of articulation, but rather how temper-

Table 2: Estimate means of closure duration predicted by the full model in *ms*, with standard error in brackets.

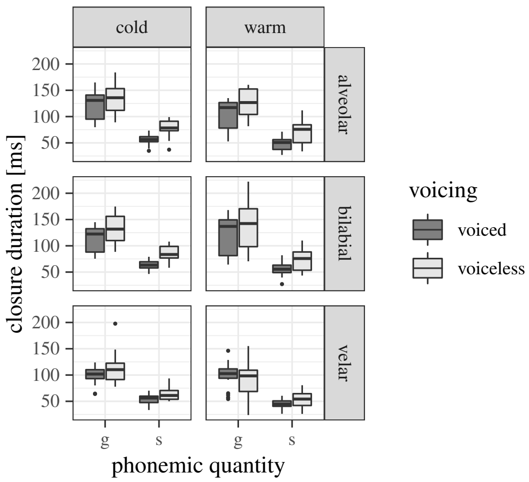
		singleton		geminate	
		cold	warm	cold	warm
alveolar	voiceless	78.8 (7.1)	69.7 (10.4)	132.5 (15.4)	126.4 (20.3)
	voiced	56.1 (6.3)	47.6 (8.7)	124.2 (13.4)	107.1 (18.2)
bilabial	voiceless	84.6 (8.5)	74.4 (12.1)	133.6 (17.1)	137.7 (21.9)
	voiced	63.1 (7.3)	57.1 (10.3)	114.7 (15.0)	123.9 (19.8)
velar	voiceless	64.5 (5.3)	53.1 (8.7)	109.9 (13.7)	92.2 (18.8)
	voiced	53.7 (4.9)	45.4 (7.0)	94.9 (11.8)	94.1 (16.8)

ature affected contrasts within one place of articulation, we decided to run separate models which would make the interpretation of interaction effects easier. For alveolars, all three main effects, QUANTITY (F[1, 4.0]= 60.6, p<0.01), VOICING (F[1, 6.5]= 19.6, p<0.01) and TEMPERATURE (F[1, 3.4]= 11.4, p<0.05), were significant, while none of the interaction effects were. In the cold condition all alveolars were produced with a longer closure duration. For bilabials, the main effects QUANTITY (F[1,3.0]=27.8, p<0.05) and VOICING (F[1,3.8]=36.5, p<0.01) were significant. While TEMPERATURE (F[1,3.0]=0.01, p=0.91) was not siginificant, the interaction of TEMPERATURE and QUANTITY was (F[1,174.8]=17.8, p<0.01). In the cold condition singletons tended to be longer compared to the warm condition while geminates were shorter. According to post-hoc comparisons the quantity difference for voiced (51($\pm$14)ms, p=0.1) and voiceless (49($\pm$14)ms, p=0.1) stops in the cold condition was not significant, while in the warm condition it was (67($\pm$14)ms and 63($\pm$14)ms, respectively). No other interactions were significant. For velars, only QUANTITY had a significant main effect (F[1,4.0]=17.6, p<0.05). There were however significant interaction effects between TEMPERATURE and VOICING (F[1,163.4]=8.1, p<0.01), as well as TEMPERATURE, VOICING and QUANTITY (F[1,163.5]=3.9, p<0.05). Voiced velars tended to be shorter than their voiceless counterparts, and velars in the cold condition tended to be longer than in the warm condition. Only voiced velar geminates in the warm condition were not shorter than the voiceless geminates in the warm condition and the voiced geminates in the cold condition.

3.2. Preceding vowel duration

We now look at the vowel duration, which we expected to be longer when preceding a geminate consonant than when preceding a singleton. The observed values are visualized in Fig. 2. While there are some differences between places of articulation, in most cases vowels preceding voiced consonants are longer than vowels preceding voiceless consonants, and vowels preceding singletons are longer than vowels preceding geminates. Statistics confirmed that QUANTITY (F[1,22.5]=5.2, p<0.05), PLACE (F[2,18.9]=21.2, p<0.01) and VOICING (F[1,23.3]=18.3, p<0.01) were significant while TEMPERATURE (F[1, 3.0]=0.0, p=0.95) was not. In addition, two-way interactions between QUANTITY and VOICING (F[1,23.5]=5.0, p<0.05) and between PLACE and VOICING (F[2,23.5], p<0.01) were significant. Model estimates are presented in Tab. 3.

Figure 1: Observed closure duration of geminate (g) and singleton (s) alveolar, bilabial and velar consonants, grouped by voicing; cold temperature condition in the left panel, warm conditon in the right panel for all three speakers.



Post-hoc analysis revealed that the quantity contrast was robust for vowels preceding voiceless but not voiced consonants. Furthermore, the voicing contrast was significant in velars but not in alveolars and bilabials.

3.3. Release duration

Release duration measurements are presented in Fig. 3 on a logarithmic scale. Following velars, release duration was overall longer for both voicing conditions, because there was some frication especially when transitioning to /i/. Release duration was also longer after voiceless stops than after voiced stops, because voiced stops were usually unaspirated. Whether the consonant was a singleton or geminate seems to not affect release duration. Statistical analysis confirmed that the main effects of PLACE (F[2,14.6]=14.8, p<0.01) and VOICING (F[1, 13.0]=30.3, p<0.01) were significant while QUANTITY (F[1,17.8]=0.05, p=0.8) and TEMPERATURE (F[1,4.0]=0.2, p=0.6) were not. Interaction effects between PLACE and VOICING (F[2,23.4]=8.4, p<0.01) and QUANTITY and TEM-

Table 3: Estimate means of preceding vowel duration predicted by the full model in *ms*, (SE).

		singleton		geminate	
		cold	warm	cold	warm
alveolar	voiceless	138.0 (12.1)	136.6 (14.5)	113.5 (11.3)	115.6 (13.4)
	voiced	145.4 (11.4)	141.3 (13.6)	144.6 (10.6)	142.5 (12.5)
bilabial	voiceless	110.6 (10.3)	111.2 (12.7)	102.4 (9.6)	102.0 (11.6)
	voiced	102.1 (9.7)	96.3 (11.9)	112.9 (9.1)	113.7 (10.9)
velar	voiceless	117.6 (10.5)	123.6 (12.9)	104.5 (9.8)	106.5 (11.8)
	voiced	149.3 (9.9)	147.3 (12.1)	128.6 (9.4)	135.3 (11.2)

Figure 2: Observed vowel duration preceding geminate (g) and singleton (s) stops.

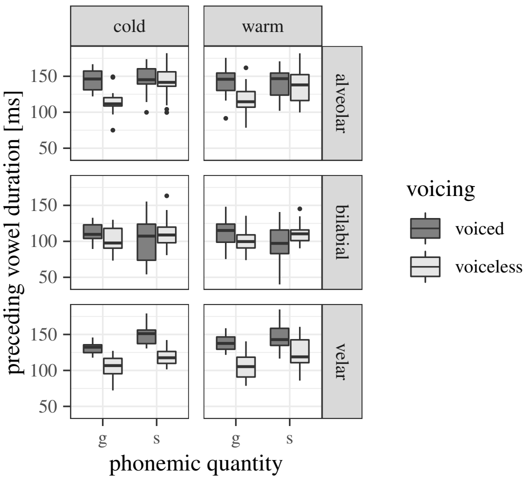
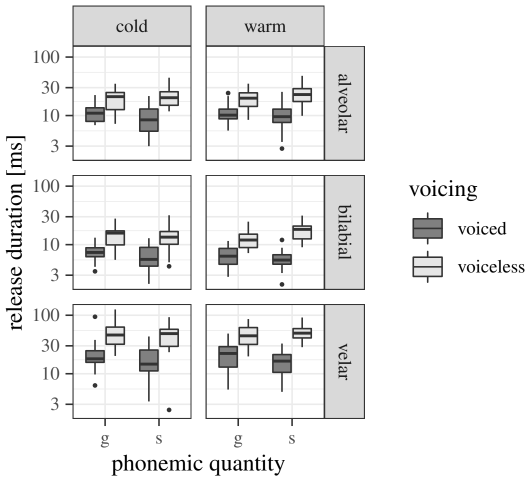


Figure 3: Observed release duration following geminate (g) and singleton (s) stops.



PERATURE (F[1,521.0]=5.5, p<0.05) were also significant. Post-hoc tests revealed that while release duration was generally longer following voiceless than voiced stops, the difference was significant only for velars. There was a tendency for release durations of singletons to be longer than of geminates in the warm condition and the other way round in the cold condition, because temperature had the opposite effect on singleton and geminate release durations (cf. Tab. 4). However, none of the differences were significant in the post-hoc comparisons.

4. SUMMARY AND DISCUSSION

First of all, we did not observe a significant main effect of temperature for any of the segments. Rather, we observed some contrast differences for the two conditions. While closure duration was generally slightly longer in the cold condition, the closure duration of bilabial geminates was shorter, so that bilabials and particularly their quantity contrast was affected by cold temperature. The release duration of geminates was longer in the cold condition than in the warm condition, but the differences were not significant and unlikely to contribute to the quantity contrast. However, our expectation that voiced segments would be especially affected by cold temperature was not met. Vowel duration was not affected at all by temperature. What we did not examine in this paper is the actual presence of vocal fold vibration during phonologically voiced segments. Often, voiceless stops were preaspirated and some vowels preceding voiced stops were devoiced or breathy towards the end. These segments were included in the vowel duration rather than stop consonant duration. Yet, it has been suggested by [19] that preaspiration contributes to consonant duration. We will thus further investigate how voicing is affected when taking into account the EGG data. It should be noted that recordings for the two temperature conditions took place in two different rooms with obvious differences in room acoustics. Furthermore, recordings took place on separate days, and with only three speakers random short-term variability of speech [15] might have contributed to some observations interpreted as a temperature effect. Comparing our data to previous studies on Italian we can say that the quantity contrast is robust and is realised by varying closure duration. When comparing the estimate means of singletons and geminates within each category, geminates are at least 1.6 times longer than singletons. This corresponds to what has been reported by [1] according to [17, p.47]. However, preceding vowel duration especially before alveolar and bilabial consonants was less affected by consonant quantity and voicing. This may have been a result of treating preaspiration as part of the preceding vowel rather than the consonant. Another possibility is that preceding vowel duration as a secondary feature for gemination is used less by South Tyrolean speakers, which might be perceived as weaker gemination.

Table 4: Estimate means of release duration predicted by the full model in *ms*, (SE).

		singleton		geminate	
		cold	warm	cold	warm
alveolar	voiceless	21.7 (6.4)	24.3 (6.1)	19.8 (7.1)	19.9 (6.6)
	voiced	7.9 (6.0)	10.3 (6.0)	12.1 (6.2)	11.8 (6.0)
bilabial	voiceless	13.6 (6.2)	17.7 (6.0)	15.0 (6.9)	12.7 (6.5)
	voiced	5.4 (5.9)	5.1 (6.0)	7.1 (6.1)	6.7 (5.9)
velar	voiceless	47.8 (8.4)	51.9 (7.8)	51.3 (9.6)	48.2 (9.0)
	voiced	17.6 (7.0)	16.6 (6.6)	21.4 (8.1)	21.0 (7.5)

5. REFERENCES

[1] Bertinetto, P. M. 1974. La quantità vocalica in italiano: verifica spettrografica. Varvaro, A., (ed), *Atti del XIV congresso internazionale di linguistica e filologia romanza* volume 3. Gaetano Macchiaroli / John Benjamins 75–85.

[2] Bertinetto, P. M., Loporcaro, M. 12 2005. The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *JIPA* 35(2), 131–151.

[3] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer. http://www.praat.org/. version 6.0.39.

[4] Esposito, A., Di Benedetto, M. G. 1999. Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America* 106(4), 2051–2062.

[5] Everett, C., Blasi, D. E., Roberts, S. G. 2015. Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences* 112(5), 1322–1327.

[6] Hansen, J. H. 1996. Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition. *Speech Communication* 20(1), 151 – 173. Speech under Stress.

[7] Kisler, T., Reichel, U., Schiel, F. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45, 326 – 347.

[8] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13), 1–26.

[9] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages*. Blackwell Publishing.

[10] Lenth, R. 2019. *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.3.3.

[11] Munroe, R. L., Fought, J. G., Macaulay, R. K. S. 2009. Warm climates and sonority classes: Not simply more vowels and fewer consonants. *Cross-Cultural Research* 43(2), 123–133.

[12] Payne, E. M. 2005. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35(2), 153–181.

[13] R Core Team, 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.

[14] Sandage, M. J., Connor, N. P., Pascoe, D. D. 2014. Vocal function and upper airway thermoregulation in five different environmental conditions. *Journal of Speech, Language, and Hearing Research* 57(1), 16–25.

[15] Sonderegger, M., Bane, M., Graff, P. 2017. The medium-term dynamics of accents on reality television. *Language* 93(3), 598–640.

[16] Steeneken, H. J., Hansen, J. H. 1999. Speech under stress conditions: overview of the effect on speech production and on system performance. *1999 IEEE International conference on Acoustics, Speech, and Signal Processing* volume 4. IEEE 2079–2082.

[17] Stevens, M. 2007. *A phonetic investigation into raddoppiamento sintattico in Sienese Italian speech*. PhD thesis University of Melbourne.

[18] Stevens, M., Hajek, J. 2010. Post-aspiration in standard Italian. *Proceedings of Interspeech 2010*. ISCA 1557–1560.

[19] Stevens, M., Reubold, U. 2014. Pre-aspiration, quantity, and sound change. *Laboratory Phonology* 5(4), 455–488.

[20] Stocks, J. M., Taylor, N. A., Tipton, M. J., Greenleaf, J. E. 2004. Human physiological responses to cold exposure. *Aviation, space, and environmental medicine* 75(5), 444–457.

[21] Vietti, A. 2017. Italian in Bozen/Bolzano: the formation of a 'new dialect'. *Towards a new standard: Theoretical and empirical studies on the restandardization of Italian* 176–212.

[22] Vietti, A. to appear. Phonological variation and change in Italian. In: Loporcaro, M., Gardani, F., (eds), *The Oxford Encyclopedia of Romance linguistics*. Oxford University Press.

[23] Westbury, J. R., Keating, P. A. 1986. On the naturalness of stop consonant voicing. *Journal of linguistics* 22(1), 145–166.

[24] Winkelmann, R., Jaensch, K., Cassidy, S., Harrington, J. *emuR: Main package of the EMU Speech Database Management System*. R package version 1.0.0.9021.