

A machine learning investigation of durational and non-durational cues to stop gemination in Italian across regional varieties and speaking rates

Angelo Dian¹ and Francesco Burroni²

¹*School of Languages and Linguistics, University of Melbourne (Australia)*, ²*Institute for Phonetics and Speech Processing, Ludwig Maximilian University of Munich (Germany)*

Introduction. Italian contrasts singleton (short) and geminate (long) consonants word-medially in its voiceless and voiced stop series /p pp t tt k kk b bb d dd g gg/, e.g., /'fato/ 'fate' vs. /'fatto/ 'fact'. This contrast is realized with considerable cross-regional phonetic differences. For instance, Northern varieties of Italian (NIt) have been claimed to produce shorter geminates compared to Central (CIIt) and Southern (SIIt) varieties, although this claim has proven controversial in light of recent experimental findings [1,2]. Importantly, no acoustic cues other than target segment duration have been explored across different regional varieties. More generally, while considerable research has focused on investigating durational acoustic correlates of Italian gemination, specifically consonant (C) and preceding-vowel (V) duration, limited attention has been given to other acoustic cues that have been found to play a role in the gemination contrast in other languages, particularly where absolute durational cues may be less reliable, e.g., at fast speaking rates (but see [3]). We address these gaps by examining different regional varieties and utilizing two distinct speech corpora that target variation in geographic region and speaking rate.

Methods. Our Corpus 1 consists of recordings of 12 speakers, six from NIt and six from CIIt. Corpus 2 involves 10 speakers, all from CIIt and SIIt. Speakers from Corpus 1 produced 16 repetitions of all Italian voiceless and voiced stop phonemes within real CVC(C)V and CVC(C)VCV words embedded in carrier phrases read at a normal pace (~2304 total tokens). Speakers from Corpus 2 produced 12 repetitions of six nonce VCV words containing target C(C) bilabial stops /ipa ippa iba ibba ima imma/ embedded in carrier phrases and read at five different speaking rates (~3600 total tokens). We test through a Random Forest Classifier (RFC, cf. [4]) the relative weight of a series of durational (dur) and non-durational (non-dur) acoustic parameters on the gemination contrast – (a) dur: C and closure duration, V duration, C/V ratio, and C/word ratio; (b) non-dur: C voicing proportion (%), f0 and RMS amplitude at the release. Accuracy we report is from 10-fold cross-validation.

Results. Results of RFC run on Corpus 1 show that dur cues dominate across varieties, with C and closure durations alone showing remarkably high accuracy (91.8% and 91.1% respectively). However, accuracy improves if non-dur cues are also included (95.8%). Slightly different patterns emerge by looking at each variety separately (cf. Table 1), with the dur cues yielding less accuracy in NIt relative to CIIt (e.g. 86.7% vs. 95.4% for C duration). The two varieties also differ in the use of non-dur cues; for instance, NIt geminates can be better classified using voicing %, while this is not the case for CIIt. Results from Corpus 2 (Table 2) show that, with variation in speaking rate, distinguishing singletons and geminates becomes a much harder task. Our best model has 83.5% accuracy. Absolute dur cues become less reliable, with performance only slightly above chance (~52% for closure duration, 64% for C duration). Better performance is achieved by relying on C/V and C/word (with both yielding ~75%) and even more so by combining relative and absolute dur cues (83.3%). Finally, no trade-off between duration and f0/RMS is observed as models based on dur and dur+non-dur cues perform similarly (83.3% vs. 83.5%).

Discussion. Our results show that different regional varieties of Italian robustly cue the contrast durationally, but the role subsumed by secondary cues, e.g., voicing %, differs by variety. With wider variation in speaking rate, we found that, rather than by absolute duration, the contrast is maintained via differences in relational durations such as C/V and C/Word (reflecting a distinct timing organization [5]). Taken together, our findings suggest a more nuanced view of phonological contrasts where the use of secondary cues is modulated by sociolinguistic factors, specifically regional variety, and contrast information is distributed beyond segments hypothesized to carry the contrast (i.e., consonants), as showcased by the better discriminative power of relational duration measures (C/V and C/Word) at faster rates.

Table 1. Results of RFC (Corpus 1)

	NIt.	CIt.
Clos. Dur.	88.98%	95.1%
C Dur.	86.7%	95.4%
C/V	78.4%	87.3%
C/Word	79.8%	89.4%
Clos. Dur. + C Dur. + C/V + C/Word + f0 + RMS	92.2%	96.8%
Clos. Dur. + C Dur. + C/V + C/Word + f0 + RMS + Voicing %	94.8%	96.7%

Table 2. Results of RFC (Corpus 2)

	CIt./SIt.
Clos. Dur.	52.1%
C Dur.	64.1%
C/V	75.2%
C/Word	74.7%
Clos. Dur. + C Dur. + C/V + C/Word	83.3%
Clos. Dur. + C Dur. + C/V + C/Word + f0 + RMS + Voicing %	83.5%

References

- [1] Mairano, P., & De Iacovo, V. (2020). Gemination in Northern versus Central and Southern varieties of Italian: A corpus-based investigation. *Language and Speech*, 63(3), 608–634.
<https://doi.org/10.1177/0023830919875481>
- [2] Dian, A., Hajek, J., & Fletcher, J. (2022). Stop (de)gemination in Veneto Italian: The role of durational correlates. In *Proceedings of the Eighteenth Australasian International Conference on Speech Science and Technology*, 176–80. Canberra, Australia.
- [3] Pickett, E. R., Blumstein, S. E., & Burton, M. W. (1999). Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56 (3–4): 135–57.
<https://doi.org/10.1159/000028448>.
- [4] Al-Tamimi, J., & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, 71, 306–325.
<https://doi.org/10.1016/j.wocn.2018.09.010>
- [5] Burrioni, F., et. al. *submitted*. Spatiotemporal features of bilabial geminate and singleton consonants in Italian.