Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

An exploratory investigation of phonological and phonetic length contrasts perception in Italian vowels and consonants

Francesco Burroni[1], *Pia Greca*[1]

[1]Institute for Phonetics and Speech Processing, LMU Munich, Germany
{francesco.burroni|greca}@phonetik.uni-muenchen.de

Abstract

Standard Italian is canonically described as a language that displays a phonological length contrast in the consonant system, but no corresponding contrast in the vowel system. Despite this fact, it is widely accepted that Italian vowels are phonetically lengthened by speakers in open stressed syllables, especially penultimate ones. However, no studies on the perception of both vowel and consonant length have been conducted. A crucial question remains open: do Italian listeners perceive the durational cues underlying a hypothesized phonological length contrast (for consonants) and a hypothesized phonetic contrast (for lengthened vowels) differently? We investigated this question in an online AX perception experiment with over a hundred Italian listeners. Results from a Mixed Effect Logistic regression model and Machine Learning classification showed that Italian listeners displayed indistinguishable identification functions for both the phonological length contrast of consonants and the "putative" phonetic durational contrast of vowels, meaning that perceptual discrimination of segmental duration was similar for phonologically long and short consonants and for vowels that were "phonetically" lengthened (or shortened) in open penultimate syllables. These results suggest therefore that Italian listeners discriminate differences in duration similarly for both consonants and vowels, either as a cue to phonological length contrasts or stress or both.

Index Terms: segmental length, Italian, perception, discrimination

1. Introduction

1.1. Length and Duration of Italian Consonants

Standard Italian is canonically described as a language that displays a phonological length contrast in the consonant system. The existence of about fifteen geminate *vs.* singleton contrastive consonant pairs is well-known from several studies [1, 2]. Consonant gemination typically occurs post-vocalically before glides, laterals, and trills, and between vowels either belonging to the same word (e.g., *fatto* [ˈfatːo] *vs. fato* [ˈfaːto]) or as a *sandhi* phenomenon between two words (known as "raddoppiamento fonosintattico", e.g., *a Roma* [aˈrːoːma]). While some Northern Italo-Romance varieties can lack gemination [2], geminate consonants belong to all regional Italian varieties from North to South [3]. The ratio between singleton and geminate duration can vary based on a variety of factors, ranging from the segment in question to the regional variety of Italian of the speaker. However, the phonological geminate/singleton contrast is maintained even at varying speech rates (slow *vs.* fast) [4]. In line with several acoustic analyses [5, 6, 7], perceptual data also confirmed that listeners are mostly sensitive to closure duration when it comes to discriminating geminate from singleton stops [8, 9], while some listeners also rely on the ratio between the consonant and the preceding vowel [10].

1.2. Length and Duration of Italian Vowels

Despite widespread length contrast for consonants, in Standard Italian (and for most varieties of the Italian peninsula with isolated exceptions [11]), there is no corresponding phonological length contrast for vowels [1, 2, 12]. However, despite a lack of phonological vowel length contrasts at the segmental level, acoustic evidence suggests that Italian vowels are phonetically lengthened by speakers in open penultimate stressed syllables, especially if compared to both word-final and antepenultimate stressed vowels [6, 13, 5, 12, 1, 14]. Given this conditioning environment, vowel lenghthening in Italian is a process which is prosodically governed at the word level [6, 2]. Additionally, in connected speech, lengthening mainly applies at the end of the intonational phrase and under emphasis [2].

According to phonological analyses [12, 1], in Italian, one can find three types of vowel lengthening, which are, however, fully predictable from the position of the vowel within a prosodic word: long vowels in penultimate stressed open syllables, semi-long vowels in stressed antepenultimate open syllables, and short vowels everywhere else. The fact that vowel length is considered as a strong acoustic predictor for stress [6] can either suggest that vowel lengthening in Italian co-occurs with stress, thus being purely phonetic, or that it is phonological and contrastive, under the assumption that only phonologically long vowels attract stress [1, 15]. However, the notion that longer durations are a pure manifestation of stress is at odds with the fact that stressed word-final vowels do not display a similar degree of lengthening, in an environment where stress and final lengthening are expected to yield an additive effect ([14] and references therein, but *cf.* [16]).

To distinguish different types of vowel lengthening, phonological analyses [12], [17, 1] have proposed that vowel lengthening is phonological in penultimate syllables but phonetic (and less salient) in antepenultimate syllables. Penultimate syllables have a preference for bimoraic feet, which arise via lengthening of penultimate syllables, given that final syllables are analyzed as extrametrical in Italian. On the other hand, lengthening in penultimate syllables is analyzed as a purely phonetic correlate of stress. Importantly, even under such phonological analyses, vowel lengthening in open penultimate syllables is not considered the reflex of contrastive segmental length, but rather a reflex of prosodically-driven lengthening.

Recent acoustic work has, however, challenged the picture of penultimate lengthened vowels being "special". According to [5], penultimate vowels are usually longer than antepenultimate ones, but the difference is not always significant. In addition, there is evidence for high variability in the duration of stressed vowels when compared to unstressed ones, both between and within-speakers and also between word types [5, 16].

10.21437/SpeechProsody.2024-91

Finally, there are no minimal pairs based on vowel length only as is the case for consonantal length.

Given this complex interplay of segmental and prosodic effects with respect to Italian vowel duration, a natural question arises: do Italian listeners make use of duration cues for vowels, and how does this compare to their use of durational cues for consonants?

Unfortunately, while there is a large body of acoustic analyses on vowel lengthening in Italian, no studies have been dedicated to the perception of vowel duration (and also relatively few have been dedicated to the perception of consonantal duration [8, 10, 9]). Perceptual work is, however, crucial to assess the role of durational cues in the segmental and suprasegmental phonology of Italian vowels. Specifically, perceptual tasks such as AX discrimination tasks have long been used to investigate the presence of both segmental length (e.g., [18]) and suprasegmental/prosodic (e.g., [19]) contrasts, thus they are ideally suited to study the status of durational differences for Italian vowels. Accordingly, in this work, we take up such a study and compare the perception of consonant and vowel duration in Italian. To preview our findings, we found no evidence of difference between the two categories, suggesting that segmental durational cues in Italian are parsed similarly between consonants and vowels, and that duration is actively recruited by Italian listeners for lexical recognition of both segmental classes.

2. Research Questions, Hypotheses, Predictions

How different durations are perceptually categorized by Italian listeners and whether this categorization is the same or different for consonant and vowels remains an open question. In other words, do Italian listeners perceive the durational cues underlying a hypothesized segmental phonological length contrast (for consonants) and a hypothesized "phonetic" prosodically-driven durational difference (for vowels) in the same way or not? Different hypotheses regarding the nature of durational contrast of Italian vowels make different predictions.

Hypothesis 1: Italian vowel durational contrasts are parsed into two phonological categories, on par with those of consonants. Alternatively, vowels may also represent a prosodic contrast that cues stress and metrical structure, i.e., a contrast that is fully integrated in the perceptual expectations of listeners, of the type suggested in [20]. **Predictions of** H_1: The prediction that stems from this hypothesis is that consontantal and vocalic durational contrast should have similar sigmoid identification functions that reflect categorical perception. No effect of contrast type (i.e., consonant *vs.* vowel) is expected on such function. Finally, we expect *not* to be able to infer above chance whether participants are listening to consonantal or vocalic stimuli based on their answer, the duration of the stimulus, and other participant/stimulus specific information.

Hypothesis 2: Alternatively, the difference between consonantal and vocalic duration is that only the former have phonological status in the grammar; while the latter represents a prosodic cue that is not parsed into phonological categories. Thus, the duration of consonants and vowels in Italian could have a function similar to pitch in cueing both tone (categorically perceived) and intonation (gradiently perceived) in the same tonal language [19]. **Predictions of** H_2: the predictions that stem from this second hypothesis are that consonantal and vocalic durational contrast should have different sigmoid identification functions that may reflect categorical perception in one case, but not in the other. Or, alternatively, two categories may exist but one is more salient in perception that the other. Effects of contrast type (i.e., consonant *vs.* vowel) on identification functions are expected. Finally, we expect to be able to infer above chance whether participants are listening to consonantal or vocalic stimuli based on their response, the duration of the stimulus, and other participant/stimulus-specific information.

3. Methodology

3.1. Participants and materials

We recruited 132 Italian listeners aged between 18-67 (μ = 36.4, σ = 12.2; 40 M, 90 F, 2 participants did not complete this answer) by advertising the experiment on social media and by means of private contacts. They took part in an AX discrimination task which was implemented using the online platform *SoSciSurvey* in its freeware version provided by the University of Munich. Participants could access the experiment via a link that was sent them via e-mail. The experiment could be run on PC or mobile devices. Only participants who declared to be L1 Italian speakers were considered for analysis, while the experiment settings allowed the automatic closure of the experiment for those who declared not to be native speakers of Italian. Due to repeated timeouts, data from 22 participants were excluded. In total we obtained 9986 usable responses.

Our stimuli were obtained from recordings of the words produced in isolation by a female native speaker of Italian trained in phonetics. These words consisted in minimal pairs (listed in Table 1). For each stimulus in the long set, we created a ten-step continuum by shortening the vowel/consonant via the removal of glottal pulses corresponding to 15 ms at each step. For the short stimuli, each vowel/consonant was stretched at each of the 10 steps by approximately 15 ms using the WSOLA method [21] implemented using Audio Toolbox in MATLAB 2023b [22]. In each trial, participants heard the original version of each stimulus and a shortened/lengthened version from the continua described above. Note that we used both long and short consonants and vowels because durational contrast can also contain secondary cues, e.g., spectral cues and f0/intensity cues that could have been lost or ambiguous if we had used only one category, *cf.* e.g., [18].

In total each participant listened to 2 (Long/Short) x 2 (Vowel/Consonant) x 3 (Words) x 10 (steps) = 120 pairs of sounds. Stimuli presentation was fully randomized for each participant. F0 and Intensity were modified to be constant (f0 flat at 175 Hz) and equal in all stimuli (average intensity normalized at 65 dB) to minimize their effects.

	Long	Gloss	**Short**	Gloss
V	amˈb**iː**to	sought	ˈamb**i**to	scope
	ruˈb**iː**no	ruby	ˈrub**i**no	(they) steal
	voˈl**aː**no	badminton	ˈvol**a**no	(they) fly
C	ˈfa**tː**o	fact	ˈfa**ː**to	fate
	ˈnɔ**nː**o	ninth	ˈnɔ**ː**no	grandfather
	ˈtu**fː**o	dive	ˈtu**ː**fo	tuff

Table 1: *Word stimuli, target segments in bold.*

3.2. Procedure

After answering a questionnaire about biographic information (geographical origin, languages spoken, age, sex, and level of education), participants were exposed to the stimuli in form of word (minimal) pairs. As displayed in Figure 1, participants

had to click on the "Play" symbol on the top of the page to listen to the word pair stimuli (each audio presented a word pair). For each stimuli pairs, participants were asked whether the two words heard were the same or different and could provide either a "Yes" or "No" answer. By clicking on "Avanti" ("Continue") participants could manually move on to the next trial.

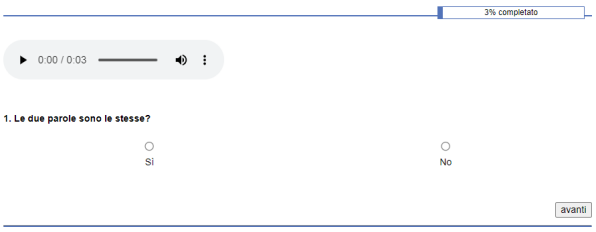

Figure 1: *Example display from the experiment platform.*

3.3. Data Processing, Independent Variables Extraction, and Statistical/Machine Learning Analyses

The results of the AX task were analyzed using Mixed Effect Logistic Regression (MELR) with the response, coded as 0 (same) and 1 (different), as the Dependent Variable (DV). The fixed effect was the continuum step z-score transformed. Note that the continuum steps range starts from -1.5 z-scores, indicating minimal modification, to 1.5 z-scores indicating maximal modification, i.e., either shortening or lengthening in 10 steps. 0, thus, represents the midpoint of the continuum. Random intercepts and slopes by Subject, Word, and Condition type (lengthening or shortening) were also added: $(DV \sim StepZ + (StepZ|SP) + (StepZ|Word) + (StepZ|Condition))$. We first analyzed separately consonants (C) and vowels (V) to determine whether their identification functions are distinct.

Additionally, we also fit a single unified model to all C and V stimuli responses $(DV \sim StepZ + Type + (StepZ|SP) + (StepZ|Word) + (StepZ|Condition) + (Type|Word) + (Type|SP) + (Type|Condition))$ to test whether the type of contrast, i.e., V or C affects the shape of the identification function. This model was compared to a null model that did not have a term for contrast type (i.e., Cs *vs.* Vs).

Finally, we reasoned that, if the behavior of Italian listeners is distinct when exposed to durational continua of C and V, then we should be able to train a Machine Learning model that can distinguish whether listeners were listening to C or V stimuli, given their answer, the participant identity, the step in the continuum, and whether the stimulus had been generated via lengthening or shortening. After experimenting with ensemble methods, Support Vector machines, and K-Nearest Neighbor (K-NN) models, we found this last class of models to perform best in the task. Thus, we employed a model that classifies Cs or Vs based on token proximity in the feature space. In this paper we report accuracy from 10-fold cross validation of a coarse K-NN model based on 100 neighbors. All statistical and machine learning analyses were performed using the Statistics and Machine Learning Toolbox in MATLAB 2023b [23].

4. Results

4.1. Separate Models for C and V duration

We fit two separate MELR models to participants responses for the C and V categories. We found that the intercept and slope have overlapping 95% CI for the two categories (V intercept: 0.15, 95% CI [-0.99 1.3]; slope: 2.9, 95% CI [2.5 3.4]; C intercept: 0.34, 95% CI [0.37 1.05]; slope: 3.27, 95% CI [2.61 3.94]). The only difference between C and V continua is represented by the wider variance in the identification function shape for the V contrasts. The MELR models' parameters result in virtually indistinguishable identification functions, as showcased by the sigmoid outputs we obtained from the models using the fixed effect coefficients, Figure 2. In particular, for both categories, the discrimination threshold lies right before the stimuli continuum midpoint, while the category transition is, for both Cs and Vs, between -0.5 and 0.5 of the continuum steps.

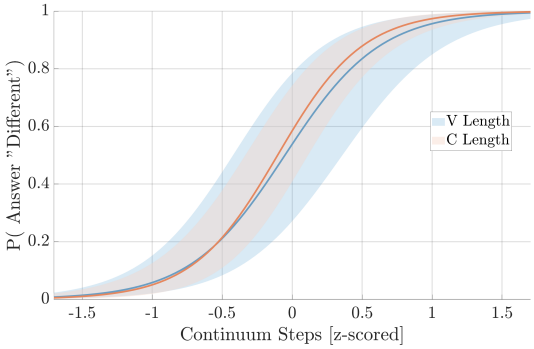

Figure 2: *Identification function obtained from MELR fit to V and C stimuli*

4.2. Unified Model for C and V duration

We fit a single MELR model to all responses belonging to the C and V categories. We used likelihood ratio test to compare the model with the fixed effect of contrast type (C *vs.* V) to a model without it. We found no evidence that the model with the additional fixed effect of $Type$ was found to be a better fit to the data ($\chi_2(\Delta df=1)$=0.13, =0.71). This was also reflected in the coefficient estimates for the contrast type, which included 0 in the alternative model (0.22, 95% CI [-0.98 1.42]), when changing from C to V stimuli. Additionally, we found an expected significant effect of step (3.08, 95% CI [2.56 3.6]), Figure 3.

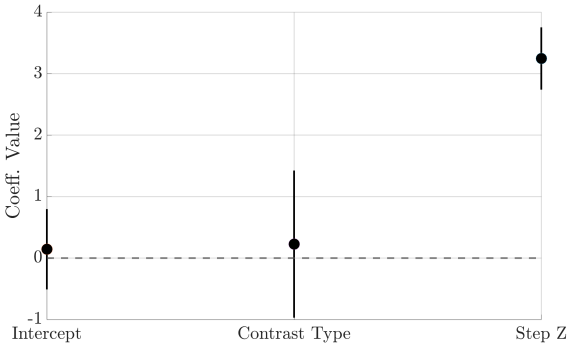

Figure 3: *Coefficient estimates of maximal model with 95% Confidence Interval, note the overlap of contrast type (C or V) with 0.*

4.3. Machine Learning Analysis

A K-NN classifier was trained to recognize whether the stimulus was a C or V, based on participant responses, participant identity, step in the continuum, and whether the stimulus had been generated via lengthening or shortening. The model

reached a micro-average accuracy of 51.8% on 10-fold cross validation. This shows that the model is basically at chance, thus corroborating the idea that C and V stimuli cannot be discriminated based on participants' behavior at a particular step in the continuum. This is in line with the findings of the statistical models. A confusion matrix of the K-NN, collapsed over all folds, is presented in Figure 4.

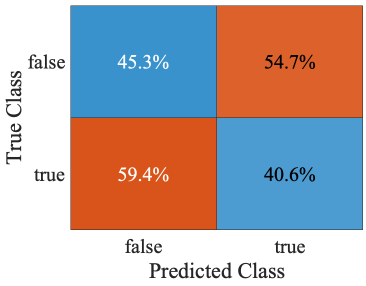

Figure 4: *Final Confusion Matrix of the K-NN classifier displaying near-chance accuracy, false = C, true = V*

5. Discussion

The purpose of our experiment was to test whether Italian listeners discriminate durational contrasts for Cs and Vs similarly. There were two hypotheses. Either Italian Vs represent a segmental contrast (like Cs) / a prosodic contrast that is fully integrated in the phonological grammar (cueing stress and metrical structure); or, alternatively, V durations represent a prosodic contrast that is not categorical but more gradient.

The results showed that Italian listeners discriminate differences in duration similarly for both segment types, either as a cue to phonological length contrasts or as a cue to prominence and metrical structure. This was showcased by the identification functions (Figure 2) and confirmed by the statistical and Machine Learning analyses graphically summarized in Figures 3 and 4. The shapes of both C and V identification functions in Figure 2 are virtually indistinguishable: they both show a discrimination threshold slightly before the midpoint and a category transition between -0.5 and 0.5 of the continuum steps, while durational variations below or above these values did not proportionally influence the responses and were all categorised similarly between Cs and Vs. Additional statistical analyses also confirmed that the main response predictor was the degree of manipulation, while the segmental class did not play a significant role in determining the type of response. Finally, a K-NN classifier trained on participant responses, participant identity, step in the continuum, and whether the stimulus had been generated via lengthening or shortening, was unable to discriminate between C and V stimuli.

For Cs, our results are in line with studies suggesting that closure duration is the main cue to gemination in stops for both Italian [8, 9] and other languages [24, 25]. Such a conclusion is also strengthened by the finding that discrimination of singleton *vs.* geminate Cs can be induced by means of manipulation of duration even in listeners who do not have this contrast in their native language [26]. As far as Vs are concerned, our results suggest that a similar identification function is obtained for V duration. V duration is also parsed into two different categories, without being, however, considered phonologically contrastive at the segmental level.

These findings are more in line with our H_1 rather than H_2. Recall that H_1 held that either Italian Vs represent a segmental contrast (like Cs) or a prosodic contrast that is fully integrated in the phonological grammar as a cue to stress and metrical structure. In our opinion, given that V duration is predictable from context and it is also a main indicator for lexical stress, as shown in previous work [6, 5, 7], the second interpretation seems more viable than the first. In other words, given its prosodic restrictions, we would not assume that similar categorical identification functions that emerged from the data point to the presence of a phonological contrast between long and short Vs for penultimate stressed syllables only. More likely, longer Vs mediate the percept of an always concomitant contrastive stress, as the two can never be separated in Italian. Yet, it is important to note that, even under the interpretation that the V categorical discrimination is due to stress changes, manipulation of the duration of Vs alone was enough to trigger a different stress percept, as f0 and intensity were controlled for. Thus, the perception of stress and prominence in penultimate syllables can (almost) entirely be reduced to duration in Italian ([6]).

A few limitations of this study should also be acknowledged. First, our data was collected in an online experiment and could therefore be more "noisy" than laboratory data. Thus, future studies are needed to confirm that the discrimination functions of Cs and Vs durations are similar also in more controlled paradigms. Second, we used an AX discrimination task to probe the perception of durational contrasts in Italian using isolated words: more work is needed to demonstrate that the findings we presented hold above the level of the prosodic word (e.g., in longer phrases and connected speech, where speech rate will be an additional factor to be considered).

6. Conclusion

In conclusion, we have presented a discrimination study of consonantal and vocalic duration in Italian. We have found that Italian listeners display near-identical identification functions for both consonants and vowels. The contrast type plays no significant role in their behavior, and a machine learning model is also unable to discriminate between consonant and vowel stimuli based on participants' behavior at different steps of the continuum. Our findings suggest that segmental durational cues in Italian – at least in isolated word pairs – are parsed similarly between consonants and vowels and in a rather categorical fashion. Duration is therefore a perceptually salient acoustic cue that is actively recruited by Italian listeners for lexical recognition. However, we find it difficult to suggest that Italian has a phonological contrast between short and long vowels similar to that of consonants. This is because of the highly restricted environments in which "long" vowels can appear. Rather, vowel duration differences may mediate the percept of prominence and prosodic structure that have categorical effects in word identification. In this respect, at least in the prosodic position and context we have studied, prominence can entirely be reduced to vowel duration. Further studies involving other syllable positions (antepenultimate and final) could help us elucidate in the future the role of vowel duration in Italian speech perception.

7. Acknowledgements

This study was supported by the European Union (ERC, *SoundAct*, project N° 101053194).

8. References

[1] M. Krämer, *The phonology of Italian*. Oxford: Oxford University Press, 2009.

[2] P. M. Bertinetto and M. Loporcaro, "The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome," *Journal of the International Phonetic Association*, vol. 35, no. 2, p. 131–151, 2005.

[3] P. Mairano and V. D. Iacovo, "Gemination in northern versus central and southern varieties of Italian: A corpus-based investigation," *Language and Speech*, vol. 63, no. 3, pp. 608–634, 2020.

[4] C. Zmarich, B. Gili-Fivela, P. Perrier, C. Savariaux, and G. Tisato, "Speech timing organization for the phonological length contrast in Italian consonants," in *Interspeech 2011 – 12th Annual Conference of the International Speech Communication Association*, 2011, pp. 401–404.

[5] S. Canalis and L. Garrapa, *Stressed vowel duration and stress placement in Italian: What paroxytones and proparoxytones have in common*. John Benjamins, 2012, p. 87–114.

[6] P. M. Bertinetto, *Strutture prosodiche dell'Italiano: Accento, quantità, sillaba, giuntura, fondamenti metrici*. Firenze: Accademia della Crusca, 1981.

[7] A. Eriksson, P. M. Bertinetto, M. Heldner, R. Nodari, and G. Lenoci, "The acoustics of lexical stress in Italian as a function of stress level and speaking style," in *Interspeech 2016, San Francisco, USA, September 8–12, 2016*. The International Speech Communication Association (ISCA), 2016, pp. 1059–1063.

[8] A. Esposito and M. G. Di Benedetto, "Acoustical and perceptual study of gemination in Italian stops," *The Journal of the Acoustical Society of America*, vol. 106, no. 4, pp. 2051–2062, 1999.

[9] L. Tagliapietra and J. M. McQueen, "What and where in speech recognition: Geminates and singletons in spoken Italian," *Journal of Memory and Language*, vol. 63, no. 3, pp. 306–323, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0749596X10000410

[10] E. R. Pickett, S. E. Blumstein, and M. W. Burton, "Effects of speaking rate on the singleton/geminate consonant contrast in Italian," *Phonetica*, vol. 56, no. 3-4, pp. 135–157, 1999. [Online]. Available: https://doi.org/10.1159/000028448

[11] D. Garassino, D. Dipino, S. Calhoun, P. Escudero, M. Tabain, and P. Warren, "Vowel length in Intemelian Ligurian: an experimental and cross-dialectal investigation," in *19th International Congress of Phonetic Sciences, Melbourne, Australia, 5 August 2019 - 9 August 2019*, 2019.

[12] M. D'Imperio and S. Rosenthall, "Phonetics and phonology of main stress in Italian," *Phonology*, vol. 16, no. 1, p. 1–28, 1999.

[13] G. Marotta, *Modelli e misure ritmiche: la durata vocalica in italiano*, ser. Fenomeni linguistici. Bologna: Zanichelli, 1985.

[14] M. Vayra, "Phonetic explanations in phonology: laryngealization as the case for glottal stops in Italian word-final stressed syllables," in *Phonologica 1992: Proceedings of the 7th International Phonology Meeting. Turin: Rosenberg & Sellier*, 1994.

[15] M. D. Saltarelli, *A phonology of Italian in a generative grammar*. University of Illinois at Urbana-Champaign, 1966.

[16] J. Hajek and M. Stevens, "Vowel duration, compression and lengthening in stressed syllables in central and southern varieties of standard Italian," in *Proceedings of Interspeech 2008*, 2008, pp. 516–519.

[17] L. D. Repetti, "The bimoraic norm of tonic syllables in Italo-Romance," Ph.D. dissertation, University of California, Los Angeles, 1989.

[18] A. S. Abramson and N. Reo, "Distinctive vowel length: duration vs. spectrum in Thai," *Journal of Phonetics*, vol. 18, no. 2, pp. 79–92, 1990.

[19] C. Gussenhoven and M. van de Ven, "Categorical perception of lexical tone contrasts and gradient perception of the statement–question intonation contrast in Zhumadian Mandarin," *Language and Cognition*, vol. 12, no. 4, pp. 614–648, 2020.

[20] J. A. Steffman, "Prosodic prominence in vowel perception and spoken language processing," Ph.D. dissertation, University of California, Los Angeles, 2020.

[21] J. Driedger and M. Müller, "A review of time-scale modification of music signals," *Applied Sciences*, vol. 6, no. 22, p. 57, 2016.

[22] T. M. Inc., "Audio toolbox version: 23.2 (r2023b)," Natick, Massachusetts, United States, 2023. [Online]. Available: https://www.mathworks.com

[23] ——, "Statistics and machine learning toolbox version: 23.2 (r2023b)," Natick, Massachusetts, United States, 2023. [Online]. Available: https://www.mathworks.com

[24] J. Hankamer, A. Lahiri, and J. Koreman, "Perception of consonant length: voiceless stops in Turkish and Bengali," *Journal of Phonetics*, vol. 17, no. 4, pp. 283–298, 1989. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0095447019304450

[25] B. B. Hansen and S. Myers, "The consonant length contrast in Persian: Production and perception," *Journal of the International Phonetic Association*, vol. 47, no. 2, p. 183–205, 2017.

[26] V. J. Porretta and B. V. Tucker, "Perception of non-native consonant length contrast: The role of attention in phonetic processing," *Second Language Research*, vol. 31, no. 2, pp. 239–265, 2015.