A comparison of rhythm metrics

in different speaking styles and in fifteen regional varieties of Italian

Rosa Giordano, Leandro D'Anna[1]

Department of Linguistic and Literary Studies, University of Salerno, Italy
`rogiordano@unisa.it, ldanna@unisa.it`

Abstract

Three different speaking styles and fifteen regional varieties of Italian are compared in order to evaluate if, and to what extent, inter-speaker variability and contextual variation can condition the temporal dynamics of speech. The analysis bases on %V, ΔC, ΔV, nPVI and rPVI metrics: as the samples of speech tested do not present internal variation in segmental phonology and in phonotactics, very similar values are expected for the three subsets. Interestingly, results show a certain range of variability across the three groups. The phenomenon is due, in our opinion, to factors external to rhythmic structure and related to high-level prosodic domains, which can also influence segmental duration patterns.

Index Terms: Rhythm Class Hypothesis, Italian, speaking styles, regional varieties.

1. Introduction

An important issue in the attempt to explain rhythmic properties of languages is represented by the Rhythm Classes Hypothesis, which resumes the long-debated question of isochrony. According to the current interpretation of the problem, rhythmic characteristics of stress-timed languages, syllable-timed languages and moraic languages would depend on different phonetic and phonological properties pertaining the segmental tier that languages belonging to the three different rhythm classes share [1,2,3,4:268]. Syllabic structure and vowel reduction are traditionally considered the most conditioning properties for rhythmic realizations, because of their correlation with segmental duration: stress-timed languages present greater variety of syllable types than syllable-timed languages and, in stress-timed languages, unstressed syllables usually have reduced vocalic systems and are consistently shorter, or even absent, with respect to stressed ones.

Several metrics, based on the properties mentioned above, have been used in order to measure duration as part of the acoustic evidence for this rhythmic dichotomy and to classify phenomena related to it [4,5,6,7]. Among them, the segmental interval measures proposed by Ramus, Nespor & Mehler [4] and the Pairwise Variability Index by Grabe & Low [5] are now largely tested on different languages. Recently, an explanation based on a more complex interpretation of the interaction of factors affecting duration in languages have been proposed by Bertinetto [7], also providing a novel testing instrument, the Control/Compensation Index.

This study aims at a comparison of different metrics on a corpus of different varieties of a single language: the measures %V, ΔC, ΔV, nPVI and rPVI are adopted, following the methods proposed in [4] and [5]. The hypothesis is that, phonology and phonotactics being equal across the subsets of data, no inconsistency in the outputs of the analysis should emerge. Expectations underlying the two metrics here used are that syllable-timed languages should exhibit [4:275] [5:527]:

- low vocalic nPVI and low intervocalic rPVI values;
- high %V values, low ΔC and possibly low ΔV values.

In current research, Italian is classified as syllable-timed language (see, e.g., [1, 3]). As for standard Italian and for regional varieties, experimental works on rhythm and on the Rhythm Classes Hypothesis generally based on limited samples of laboratory speech or read speech elicited in controlled conditions ([1,3] for a review, [8]); few studies investigated the temporal dynamics in spontaneous speech [9,10,11,12,13] and in different speaking styles [13]. Recent works testing %V, ΔC and ΔV [4, 14, 15] or nPVI and rPVI indices [16] also based on read speech. Articulatory properties of prominence and word-level coarticulation indices have been analysed as well (among other works, [17]).

Regional varieties, which are also considered here (read and dialogic *corpora*), do not differ from Standard Italian (TV *corpus*) in their segmental phonology or phonotactics [18]; they can present differences in the phonetic realization of segmental phonology and vowel and consonant lengths can be affected in some cases, with no clear systemic regularities.

Italian dialects, which are not the object of the present work, are instead the original Italo-romance languages and actually differ from Standard Italian both in phonological inventories and syllabic structures. The study by Schmid [19] concerned Italian dialects and tested duration patterns basing on %V, ΔC and ΔV metrics applied to spontaneous speech.

Previous works on read [4, 14, 15, 16] or spontaneous Italian [9, 10, 11, 12, 13], although generally confirmed predictions regarding syllable-timed language, highlighted a certain degree of variability of the values related to these metrics, which in some cases resulted similar to those expected for isoaccentual languages. Results in [19] prove the general tendency for Italian dialects to reach higher %V (%V>50) than standard Italian or Italian regional varieties.

2. Corpus and analysis

2.1. Corpus

The analysis bases on a *corpus* of Italian of more than 12000 syllables, containing samples of speech from 34 speakers, divided in three sets of data manually segmented and labeled.

1) Pre-planned monologic speech: four speech streams selected from four TV News of the main national Italian TV-channels (Standard Italian; 4 journalists, professional speakers; 1070 syllables).

2) Spontaneous dialogic speech: two task-oriented dialogues selected from the Italian national *corpus* CLIPS (regional varieties: Rome and Perugia; 4 speakers; about 5000 syllables).

3) Read speech: ten sentences selected from the Italian national *corpus* CLIPS and read by thirty speakers (regional varieties; 5940 syllables).

The CLIPS *corpus* consists of fifteen sets of spontaneous and read speech representing as many regional varieties of Italian. Speakers are homogeneous as for sociolinguistic criteria (university students, aged 25-30, born and living in the town they represent) [20,21].

Table 1. *Vocalic and intervocalic intervals per subset.*

	Vocalic intervals	Intervocalic intervals	Total
Read	4119	4674	8793
Dialogic	5218	6270	11488
Pre-planned	1030	938	1968
Total number	*10367*	*11882*	*22249*

2.2. Metrics

Following the methods illustrated in [4] and [5], which did not require significance analysis, we calculated the following indices for all the vocalic and the intervocalic sequences in a sentence and/or a prosodic group, except for hesitations, interruptions or sequences not segmented because of their phonetic reduction or underspecification (Table 1).

%V is the proportion of vocalic intervals within the sentence, that is the sum of vocalic intervals divided by the total duration of the sentence; it is related to syllabic structure.

ΔC, related to syllabic structure, is the standard deviation of the duration of intervocalic intervals within each sentence.

ΔV is the standard deviation of the duration of vocalic intervals within each sentence, which generally is not considered as indicative as %V and ΔC for the motivation listed in Ramus et al. [4: 274-5].

rPVI, row Pairwise Variability Index - Vocalic, expresses the level of variability across successive measurements for vocalic intervals [5].

nPVI, normalised PVI – Intervocalic, is compiled by dividing the difference in duration between each pair of successive measurements by the mean duration of the pair; differences are then summed and divided by the number of differences [5]; speech rate effects should be avoided.

3.   Results and discussion

3.1. Speaking styles

Table 2 and graphs 1 to 3 show results.
In dialogic and read speech, %V values are consistent with the results expected for syllable-timed languages, while pre-planned speech shows values similar to those expected for moraic languages. ΔV and ΔC are generally consistent with the results expected for syllable-timed languages, although standard deviations strongly vary, as follows:

	READ	DIALOGIC	PRE-PLANNED
ΔC	-	→	+
ΔV	+	←	-

As far as PVI indices are concerned, normalized values are consistent with the results expected for isosyllabic languages while Intervocalic rPVI is near to what registered for ΔC.

Table 2. *Mean values for intervocalic rPVI, vocalic nPVI, %V, ΔV, ΔC and standard deviations (Δ or SD) for each subset of data.*

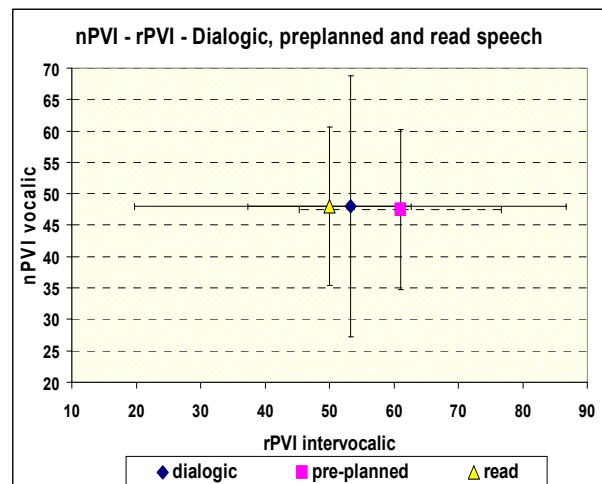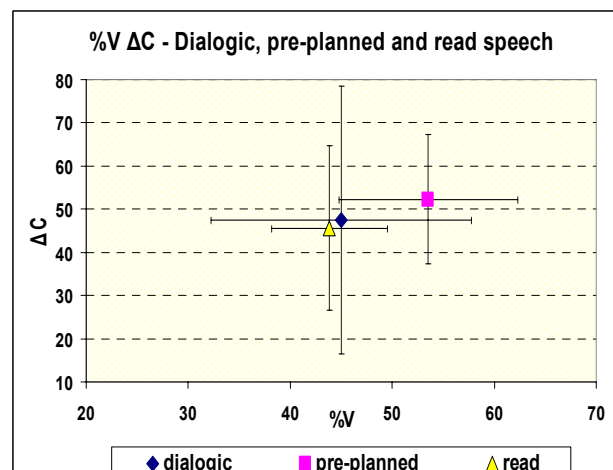| Metrics	Pre-planned	Read	Dialogic
Int. rPVI	61.0	50.0	53.2
Δ rPVI	*15.7*	*12.7*	*33.5*
Voc. nPVI	47.5	48.0	48.1
Δ nPVI	*12.7*	*12.6*	*20.8*
%V	53.5	43.9	45.0
SD %V	*8.8*	*5.7*	*12.8*
ΔV	40.7	46.5	42.8
SD ΔC	*15.0*	*19.0*	*31.1*
ΔC	52.3	45.6	47.5
SD ΔV	*10.8*	*12.3*	*32.8*

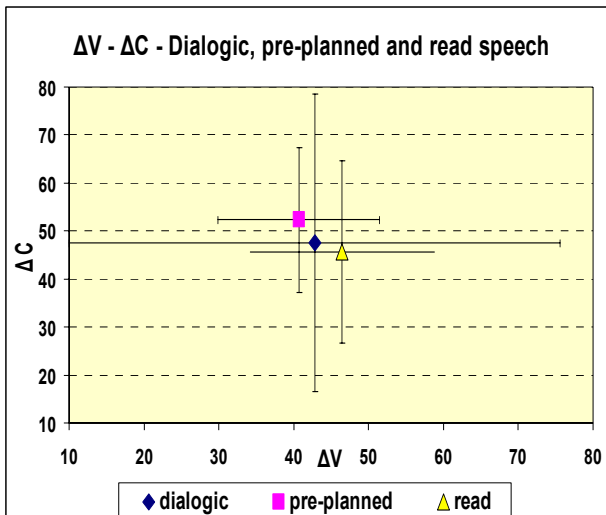
Figure 1: *nPVI and rPVI ratio.*

Figure 2: *%V – ΔC ratio.*
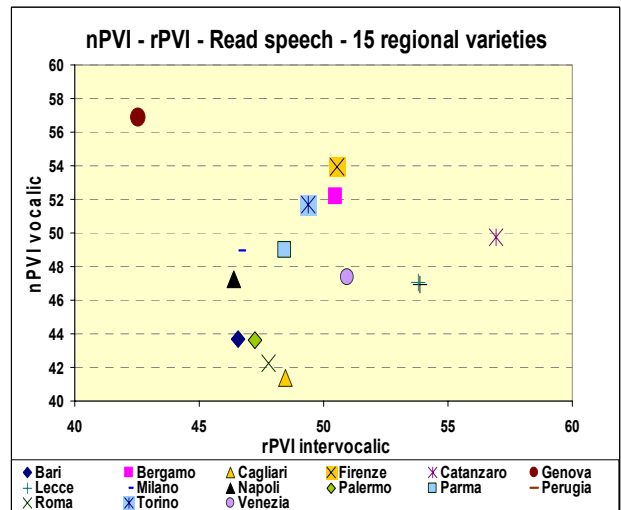
Figure 3: *ΔV – ΔC ratio*.

Figure 4: *nPVI and rPVI ratio*.

3.2. Regional varieties

In regional varieties, %V values are generally consistent with isosyllabic languages and are clearly different from dialects (Schmid 2004: %V > 50), although in several cases they are near to what is expected for isoaccentual languages (%V: 40-42). ΔV and ΔC values also show great variability: in Perugia and Catanzaro Italian ΔV is near to isoaccentual languages while in Naples Italian ΔC is similar to moraic languages. nPVI and rPVI also show similar oscillation.

Table 4. *Mean values for intervocalic rPVI, vocalic nPVI and their standard deviations*.

	Int. rPVI	Δ rPVI	Voc. nPVI	Δ nPVI
Bari	46.6	14.1	43.7	9.1
Bergamo	50.5	12.6	52.2	13.1
Cagliari	48.5	11.5	41.4	12.6
Catanzaro	56.9	15.1	49.7	12.3
Firenze	50.6	12.4	53.9	13.2
Genova	42.5	13.0	56.9	15.8
Lecce	53.8	14.9	47.0	11.5
Milano	46.6	11.8	48.9	14.4
Napoli	46.4	10.5	47.3	14.4
Palermo	47.3	12.7	43.6	14.3
Parma	48.4	11.8	49.0	11.8
Perugia	63.9	15.7	45.5	9.1
Roma	47.8	12.7	42.2	10.7
Torino	49.4	11.6	51.7	12.9
Venezia	50.9	9.6	47.4	13.1

Table 5. *Mean values for %V, ΔV, ΔC and their standard deviations*.

|	%V	Δ%V	ΔV	SD ΔV	ΔC	SD ΔC
Bari	45.5	5.3	45.4	15.3	50.5	19.0
Bergamo	45.5	5.3	46.2	11.3	46.2	18.4
Cagliari	43.0	6.7	45.1	9.8	41.3	21.1
Catanzaro	42.3	5.4	54.7	14.1	55.8	25.4
Firenze	46.4	5.9	47.0	15.7	44.2	18.9
Genova	42.3	6.2	41.0	11.8	45.9	19.0
Lecce	43.1	5.6	48.1	12.3	46.7	16.3
Milano	45.5	6.4	43.4	11.1	41.3	17.8
Napoli	41.0	4.6	43.2	9.6	35.8	18.3
Palermo	42.8	6.9	45.0	14.6	41.8	22.8
Parma	45.8	6.1	45.3	11.5	43.2	19.5
Perugia	48.4	4.9	56.2	13.4	55.0	19.8
Roma	44.3	5.0	43.7	14.2	48.6	18.2
Torino	40.9	5.6	45.2	10.0	44.3	14.6
Venezia	41.0	5.3	48.1	10.1	43.1	16.2
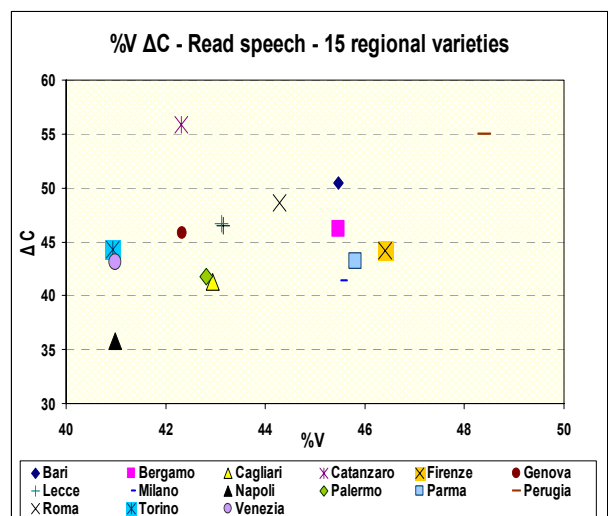
Figure 5: *%V – ΔC ratio*.
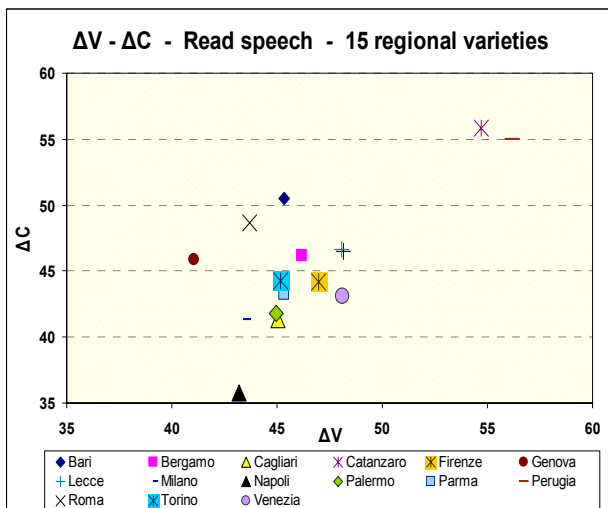
Figure 6: *ΔV – ΔC ratio*.

3.3. Discussion

Both for diamesic and diatopic varieties of Italian, high degree of variability of segmental durations is evident in measures which have not been normalized. The point is very important, because while phonological properties usually related to the classification for the Rhythm Class Hypothesis do not change, rhythmic indices outputs vary from values consistent with syllable-timed languages to values usually given for stress-timed languages or, in some cases, for moraic languages. Such results would weaken the predictive power of such metrics and bring us to try to explain causes and consequences of such intra-linguistic variability.

Previous works on the same materials show that the variability of segmental durations here registered can be preliminarily explained as related to both speech rate [12] and to speaking styles [13]: in particular, TV journalists generally tend to hyperarticulation and frequently use a particular degree of emphasis, which can also be signaled by non-pitch features in several languages and which can affect vowel durations in the material analyzed. Moreover, several prosodic factors can affect duration variability [7,8,13]. A temporal dynamics corresponding to the rhythmical pattern actually occurs in the speech flow: the majority of stressed (i. e. rhythmically strong) syllables is longer than the preceding unstressed (i.e. rhythmically weak) syllable. Apart from rhythmic factors, vowel duration is related to the intonation level, as unstressed syllables always present lower values than accented syllables for all the speakers, and accented syllables also tend to be longer than stressed syllables [13].

4. Conclusion

This work provides evidence for a wide-ranging variation of duration values in different speaking styles and in different regional varieties of the Italian language. The output of the metrics used are not completely consistent with the hypotheses of output underlying the rhythmical representation in [4] and [5]. The point to be better investigated is probably the interplay between the rhythmic structure and the intonation tier of a language, which can also influence, although indirectly, duration patterns (see also [7][12][13]). This could lead to improve our capability to explain and capture rhythmic regularities, as also argued in [22].

5. References

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

[2] Dauer, R., "Stress-timing and syllable-timing Reanalyzed", Journal of Phonetics 11:51-62, 1983.

[3] Bertinetto, P. M., "Reflections on the dichotomy 'stress vs. syllable-timing'", Revue de Phonétique Appliquée, 91:99-130, 1989.

[4] Ramus, F., Nespor, M., Mehler, J., "Correlates of linguistic rhythm in the speech signal", Cognition, 73(3):265-292, 1999.

[5] Grabe, E., Low, E.L., "Durational variability in speech and the Rhythm Class Hypothesis", in Gussenhoven C., Warner N. [eds.], Laboratory Phonology 7, 515-543, Berlin: Mouton de Gruyter, 2002.

[6] Gibbon, D., Gut U., "Measuring speech rhythm", Proceedings of Eurospeech 2001, Aalborg, Kommunik Grafiske Losninger A/S, 95-98, 2001.

[7] Bertinetto, P. M., Bertini, C., "On modeling rhythm in natural languages", Proceedings of the 4th Speech Prosody Conference, Campinas BR, 2008.

[8] Marotta, G., Aspetti della struttura ritmico-temporale dell'italiano, ETS, Pisa, 1984.

[9] Crocco, C., Giordano, R., "Un'indagine sull'isocronia nell'italiano parlato connesso", Atti delle X Giornate di Studio del Gruppo di Fonetica Sperimentale, 59-67, 2000.

[10] Barry, W., Russo, M., "Isocronia soggettiva o oggettiva? Relazioni tra tempo articolatorio e quantificazione ritmica", in Albano Leoni F., Cutugno F., Pettorino M., Savy R. [eds.], Il parlato italiano, Napoli: D'Auria, CD-Rom, (A02), 2004.

[11] Gamal, D., "Sul ritmo in italiano L2: l'eventuale isocronia". In Savy R., Crocco C. (eds.), Atti del II Convegno nazionale dell'AISV. CD-rom. Torriana: EDK Editore, (2006). Online at www.parlaritaliano.it

[12] Giordano, R., "Note sulla fonetica del ritmo dell'italiano", in Savy R., Crocco C. [eds.], Atti del II Convegno Nazionale AISV, Torriana: EDK Editore, CD ROM, 2006. Online at www.parlaritaliano.it

[13] Giordano, R., "On the phonetics of rhythm of Italian: patterns of duration in pre-planned and spontaneous speech", Proceedings of the 4th Speech Prosody Conference, Campinas, BR, 2008.

[14] Mairano, P., Romano, A., "Inter-subject agreement in rhythm evaluation for four languages (English, French, German, Italian)", Proceedings of the XVIth ICPhS, Saarbrücken, 1149-1152, 2007.

[15] Dellwo, V., Fourcin, A., Abberton, E., "Rhythmical classification of languages based on voice parameters", Proceedings of the XVIth ICPhS, Saarbrücken, 1129-1132, 2007.

[16] Dellwo, V., Steiner I., Aschenberner B., Dankovičová J., Wagner P., "BonnTempo Corpus and BonnTempo Tools: A database for the study of speech rhythm and rate", Proceedings of the 8th ICSLP, Jeju, Korea, 2004.

[17] Avesani, C., Vayra, M., Zmarich, C., "On the articulatory bases of prominence in Italian", Proceedings of the XVIth ICPhS, Saarbrücken, 981-984, 2007.

[18] Bertinetto P.M., Loporcaro M., "The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome", IPA Journal, 35(2):131-151, 2005.

[19] Schmid, S., "Une approche phonétique de l'isochronie dans quelques dialects italo-romans", in Meisenburg T., Selig M. [eds.], Noveaux départs en phonologie. Les conceptions sub- et suprasegmentales, Tübingen: Narr, 109-124, 2004.

[20] Albano Leoni, F., Cutugno, F., Savy, R. [eds.], CLIPS corpus (Corpora e Lessici di Italiano Parlato e Scritto). 2005. Online at http://www.clips.unina.it

[21] Savy, R., Cutugno, F., "CLIPS: diatopic, diamesic and diaphasic variations of spoken Italian", Proceedings of the 5th Corpus Linguistics Conference, Liverpool, 2009.

[22] Kohler K.J. (ed.), Phonetica, Thematic Issue on Rhythm in Speech and Language, 66(1-2), 2009.