Comparison of MSA Rhythm metrics in three southern Algerian regions

Droua-Hamdani Ghania

الملخّــص

يتناول هذا المقال حساب مقاييس الإيقاع للّغة العربيّة الفصحى (MSA) اعتمادا على تسجيلات صوتيّة لمناطق الجنوب الجزائريّ (بشّار وغرداية والوادي). وقد أُخذت هذه المدوّنة من قاعدة البيانات ALGASD. أجريت التّجارب على الإيقاع باستعمال الأمثلة التّالية : IM و PVI وCCI وفقا للمنشإ (المنطقة) وجنس المتكلّمين. أمّا التّجربة الأخيرة فتعلّقت بمقارنة بين قيم الإيقاع المتحصّل عليها في التّجارب الأولى والنّتائج التي توصّلنا إليها باستخدام قاعدة البيانات West Point للّغة العربيّة.

الكلمات المفاتيح : مقاييس الإيقاع، الأمثلة IM وPVI وCCI، اللّغة العربيّة الفصحى، المناطق، وجنس المتكلم.

Comparison of MSA Rhythm metrics in three southern Algerian regions

Droua-Hamdani Ghania

Scientific and Technical Research Center for the Development of Arabic Language, Algiers, Algeria.

gh.droua@post.com

Abstract

This paper deals with rhythm metrics of MSA language computed from speech files of three southern regions of Algeria (Bechar, Ghardaia and El Oued). The data collection is a part of ALGASD corpus. The experiments were conducted to investigate the variation of IM, PVI and CCI scores according to origin and gender of speakers. The last investigation concerns the comparison of these rhythm values with results finding for West Point corpus.

Key words: Rhythm metrics, IM, PVI and CCI models, MSA, regions, gender.

1. Introduction

Many studies have put forward a set of metrics to reveal rhythm differences and similarities between languages. Examples of experimentation addressed on languages: English, French, Italian, Spanish, etc. [1-4]. These metrics have been used also to carry out second language acquisition characteristics, by focusing on the impact of the first language (L1) on the rhythm of the second language (L2) such as: – Korean English; [5] – Singapore English; [6] – German L2 influenced by Chinese, English, French, Italian and Romanian L1; [7]. Likewise studies on Arabic rhythm were conducted these last years as: the comparison of MSA with other languages –stressed/timed languages- [8-9], comparison of MSA with Arabic dialects [10], the impact of rhythm L1 (English) on L2 (MSA) acquisition [11].

Speech rhythm is based on the acoustic durations of vocalic and consonantal intervals of the vocal signal. The algorithms performed for this purpose are: Interval Measures (IM) and Pairwise Variability Indices (PVI). The IM approach involves computing three separate measures from the segmentation of speech signals (global utterance) into vocalic (V) and consonantal (C) units (ΔV, ΔC and $\%V$) [12]. The time-normalized metric measures (VarcoV/C) were introduced when it was observed that the consonantal interval measure is inversely proportional to speech rate [13]. The PVI (Pairwise Variability Index) algorithm differs from the IM and the VarcoV/C models in that it focuses on the temporal succession of the vocalic and consonantal intervals instead of global utterance [14]. The model suggests that the rPVI should be used for the consonantal intervals, while the nPVI (normalized Pairwise Variability Index) should be used for the vocalic intervals. The formulas to compute nPVI and rPVI are respectively:

$$nPVI = 100 \times \frac{\left(\sum_{k=1}^{m-1} \left| \frac{(d_k - d_{k+1})}{(d_k + d_{k+1})/2} \right| \right)}{(m-1)} \qquad (1)$$

$$rPVI = \frac{\left(\sum_{k=1}^{m-1} \left| d_k - d_{k+1} \right| \right)}{(m-1)} \qquad (2)$$

where m is the number of intervals, and d is the duration of the k^{th} interval.

The third model proposed to compute rhythm of languages has been developed by [15]. Comparing to the first models, the latest is based on syllable complexity. This algorithm, which is called Compensation and Control Index (CCI), is inspired by the syllable compensation. CCI model states that controlling languages (syllable-timed) are supposed to show low levels of compensation at intra and inter-syllabic in comparison to compensating languages (stressed-timed) that are supposed to show higher levels of compensation. So, the CCI metric has for task computing the level of compression (lengthening or shortening) acceptable in a language according to the context. CCI formula is given by:

$$cci = \frac{100}{(m-1)} \sum_{k-1}^{m-1} \left| \frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}} \right| \qquad (3)$$

where m is the number of intervals, d for duration, and n for number of segments within the relevant interval.

CCI algorithm assumes therefore that the controlling languages should present similar C and V local durational variations dotted along the bisecting line. However, the compensating languages should vary more in the vocalic than in the consonantal segments; they should be gathered under the bisector [15] (figure1).

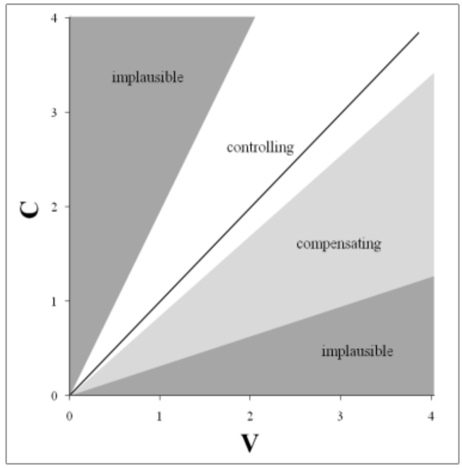

Figure 1: Schematic representation of the major rhythmic types according to the CCI model [15].

The experiment dealt with Modern Standard Arabic language (MSA) produced by Algerian speakers. Regional linguistic variations of MSA in Algeria have been the focus of only a very small number of works especially in rhythm (prosodic field).

This study assessed the rhythm properties of MSA language spoken in three regions scattered across the south of the country. The regions are: Bechar in the west (close to Morocco), Ghardaia in the center of Algeria' south and El Oued is in the East (close to Tunisia). The rhythm metrics that are examined: three interval measures (%V, ΔV, and ΔC), two time-normalized indices (VarcoV, VarcoC), and two pairwise variability indices (nPVI-V, rPVI-C), and finally two compensation and control indices (CCI-V, CCI-C).

The paper is organized as follows: Section 2 gives the main specificities of the MSA language. Section 3 presents an overview of the speech material and speakers used for the purpose of the rhythm studies. Section 4 carries out the results and discussions about different rhythm experiments on MSA of three southern regions in Algeria. Section 5 concludes this work.

2. MSA Background

Modern Standard Arabic (MSA) uses a distinct phonological system that includes 28 consonants and six vowels: three short vowels vs. three long vowels.

There are two distinctive classes of MSA consonants, pharyngeal and emphatic consonants. In addition to these classes, Arabic is characterized by two distinctive features that are fundamental to avoid ambiguity in word meanings: long vowels and gemination. The latest characteristic or consonant elongation happens when a spoken consonant is pronounced for a perceptibly longer period of time than a short consonant. All Arabic consonants may be geminated.

Arabic language is endowed by two kinds of syllables: open syllables (CV) and (CV:)— and closed syllables (CVC), (CV:C) and (CVCC), V indicates a vowel while C indicates a consonant. The Arabic vowels never occur initially. All Arabic syllables must contain at least one vowel.

3. Data collection and speakers

Speech material used for the purpose was taken from ALGASD corpus [16] which consists of collection of 200 balanced sentences recorded from 300 speakers from 11 regions across Algeria. The database consists of 1080 wave files. Recordings were captured at a sampling rate of 16 bit at 16 kHz.

The present study used only a set of recording relating to southern regions: Bechar in the west (R1), El Oued (R2) and Ghardaia (R3). To reduce factors that can compromise the reliability of rhythm analysis, a large sample of speakers was used and the measurement effects across speech materials were controlled for. So, speech material from 37 speakers (19 males/ 18 females) reading 136 recordings (three to six sentences per speaker) were used. Table 1 shows the number and gender of speakers in the sample.

Regions	male	female	Total
R1	4	3	7
R2	8	8	16
R3	7	7	14

Table 1. Distribution of speakers in southern regions

The second MSA speech material used in this paper is the West Point corpus which is collected and processed by at the United States Military Academy at West Point and the Center for Technology-Enhanced Language Learning [17]. The corpus was conceived to train acoustic models for automatic speech recognition. For the purpose of this work, a speech material of 15 Arabic speakers (10 female and 5 male) reading 75 recordings (five sentences per speaker) were used.

All vowels and consonants of both corpora were segmented by inspection of speech waveforms and wideband spectrograms by one measurer. Vowel and consonant durations were extracted using a customized script on the boundary label files. Rhythm metrics were computed for each sentence and for each speaker.

ments on MSA of three southern regions in Algeria. Section 5 concludes this work.

2. MSA Background

Modern Standard Arabic (MSA) uses a distinct phonological system that includes 28 consonants and six vowels: three short vowels vs. three long vowels.

There are two distinctive classes of MSA consonants, pharyngeal and emphatic consonants. In addition to these classes, Arabic is characterized by two distinctive features that are fundamental to avoid ambiguity in word meanings: long vowels and gemination. The latest characteristic or consonant elongation happens when a spoken consonant is pronounced for a perceptibly longer period of time than a short consonant. All Arabic consonants may be geminated.

Arabic language is endowed by two kinds of syllables: open syllables (CV) and (CV:)— and closed syllables (CVC), (CV:C) and (CVCC), V indicates a vowel while C indicates a consonant. The Arabic vowels never occur initially. All Arabic syllables must contain at least one vowel.

3. Data collection and speakers

Speech material used for the purpose was taken from ALGASD corpus [16] which consists of collection of 200 balanced sentences recorded from 300 speakers from 11 regions across Algeria. The database consists of 1080 wave files. Recordings were captured at a sampling rate of 16 bit at 16 kHz.

The present study used only a set of recording relating to southern regions: Bechar in the west (R1), El Oued (R2) and Ghardaia (R3). To reduce factors that can compromise the reliability of rhythm analysis, a large sample of speakers was used and the measurement effects across speech materials were controlled for. So, speech material from 37 speakers (19 males/ 18 females) reading 136 recordings (three to six sentences per speaker) were used. Table 1 shows the number and gender of speakers in the sample.

Regions	male	female	Total
R1	4	3	7
R2	8	8	16
R3	7	7	14

Table 1. Distribution of speakers in southern regions

The second MSA speech material used in this paper is the West Point corpus which is collected and processed by at the United States Military Academy at West Point and the Center for Technology-Enhanced Language Learning [17]. The corpus was conceived to train acoustic models for automatic speech recognition. For the purpose of this work, a speech material of 15 Arabic speakers (10 female and 5 male) reading 75 recordings (five sentences per speaker) were used.

All vowels and consonants of both corpora were segmented by inspection of speech waveforms and wideband spectrograms by one measurer. Vowel and consonant durations were extracted using a customized script on the boundary label files. Rhythm metrics were computed for each sentence and for each speaker.

rics were measured (vocalic /consonantal compensation and control Index – CCI-V/CCI-C respectively–) for each sentence of MSA R1, R2 and R3 speakers. Table 3 shows the average values of each CCI-C and CCI-V computed for the three regions.

Regions	**CCI-V**	**CCI-C**
R1	154.17	161.95
R2	132.39	174.15
R3	139.71	144.90

Table 3. CCI measures by regions

The comparison of the vocalic/consonantal compensation and control Index metrics reveals that R1 shows a higher CCI-V value compared to R2 and R3 measures whereas, CCI-C score is raised for R2 rather than R1 and R3. When comparing the pair (CCI-V, CCI-C) valuies between all regions, it can be seen that R2 measures are less great than R1 and R3.
These findings put forward that the reduction of vowel duration produced by R2 speakers is compensated by elongations of consonant units (quantity). So, the syllable is sustained through the consonant compensation. Then, MSA R2 seems to be more compensating language than controlling. Indeed, the vocalic units vary more than consonantal segments. In contrast, R1 and R3, show a similar tendency of C and V local durational variations. They assume to be controlling languages.

4.3. IM, PVI and CCI measures according to speakers' gender

To reveal possible differences or similarities in rhythm patterns of speakers according to their gender, an examination of the acoustic characteristics of MSA vowels and consonants depending on that factor is conducted. Thus, all seven metrics were computed considering both kinds of speakers (females vs. males). Figures 3 and 4 illustrate fluctuations of rhythm metrics according to speakers' gender and region.
As it can be seen from figure 3 and 4, all rhythm metrics of R3 female speakers are lower than R3 male values. The findings show also, that R2 females' scores are similar to R2 males' measures except for PVI values where females' scores are higher than PVI of males. Regards to R1 values, the %V, ΔV and rPVI for female speakers are great than their males' counterparts. In contrast, (ΔC, VarcoV and VarcoC) of females are lower than males values. nPVIs of both gender is close to each others. The plots illustrate also that globally, rhythm measures of R1 females are higher than R2 ones followed by R3 scores. However, in the case of males' speakers, R2 values are lower than R3 and R1 respectively.
One-way ANOVA was performed to test the effect of the gender of speakers on vowel duration values. The results show a highly significant effect for VarcoV and ΔV ($F_{(1,130)}=7.65$, $p=.006$ and $F_{(1,130)}=3.45$, $p=.06$ respectively).

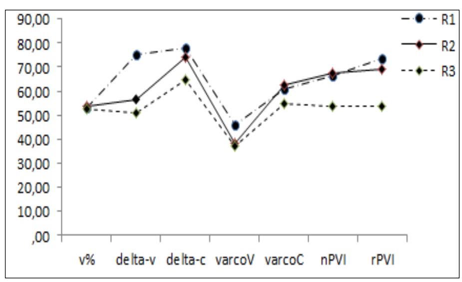

Figure 3: Comparison of rhythm metrics between MSA females

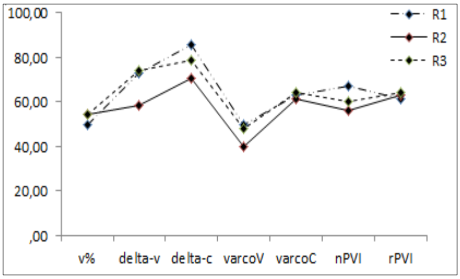

Figure 4: Comparison of rhythm metrics between MSA males

The outcomes of this experiment expose that the vocalic rhythm measures depend on the gender of the speakers. Indeed, a reduction of vowel duration is observed when speakers are females in comparison with the production of vowels by males.

4.4. MSA Algerian southern regions vs. West Point MSA

The last experiment consists on a comparison of the mean values of rhythm metrics computed across all southern regions (R1, R2 and R3) with MSA West Point rhythms values finding in [11]. Table 4 reports mean values of each of the nine rhythm measures.

Metrics (ms)	Southern Regions	West Point
%V	52.99	42.41
ΔV	64.73	49.42
ΔC	75.26	53.29
VarcoV	43.13	65.14
VarcoC	61.21	50.81
rPVI-C	61.81	74.84
nPVI-V	64.41	56.22

Table 3. Comparison of rhythm values between West Point and Algerian southern regions MSA

As it can be seen, all vocalic measures assessed in the southern regions are higher than West Point measures except for vocalic time-normalized VarcoV. Likewise, ΔC, VarcoC and rPVI-C produced by southern regions speakers are great than their counterparts of West Point. These unexpected findings seem showing that southern speakers produce more elongation in duration of vocalic and consonant segments than West Point speakers for those corpora.

5. Conclusion

This study focused on the rhythm properties of MSA language spoken in three regions scattered across the south of the Algeria. Thus, the experiments deal with the rhythmic features across MSA language produced by speakers of three regions in the south (Bechar -R1-, Ghardaia -R2- and El Oued -R3-). The outcomes show that speakers of R2 present several differences in rhythm scores compared to R1 and R3 values especially in vocalic segments. These rhythm variations are performed using three rhythm algorithms the IM, PVI and CCI approaches for both vowel and consonant scores. In light of the statistical findings, vocalic metrics show significant effects of speakers' regions origin and gender on vowels' duration. According to these results, it thus seems that R2 speakers (El Oued -close to Tunisia-) produce vowels durations that are reduced comparing with their counterparts of R1 and R2. The last experiment consist on a comparison of the mean values of rhythm metrics computed across all southern regions with those finding for West Point MSA. Results reveal that southern speakers produce vowels and consonants that are more stretched in term of duration comparing to those assessed in West Point corpora.

The main aim of the study is to search how revealing the regional linguistic variations of MSA in Algeria which have been unfortunately studied only by a very small number of works. These results will be used in future works to improve an ASR system realized using Algerian speech corpus namely ALGASD.

6. References

[1] Arvaniti, A., “Rhythm timing and the timing of rhythm”, Phonetica 66, 46-63. 2009.

[2] Nolan, F. and Asu, E., “The pairwise variability index and coexisting rhythms in language, Phonetica 66, 64-77, 2009.

[3] Mairano, P. and Romano, A., “Rhythm metrics for 21 languages”, Proceedings of ICPhS XVII, Hong Kong, 17–21, 2011.

[4] White, L. and S.L. Mattys, “Rhythmic typology and variation in first and second languages”, Segmental and Prosodic issues in Romance Phonology. pp. 237-257. Amsterdam: John Benjamins, 2007.

[5] Jang, T.Y., “Automatic assessment of non-native prosody using rhythm metrics: Focusing on Korean speakers' English pronunciation”, In Proc. of the 2nd International Conference on East Asian Linguistics, 2009.

[6] Ling, L-E., Grabe, E. and Nolan, F., “Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English”, Language and speech, 43(4):377–401. 2000.

[7] Gut, U., “Non-native speech rhythm in German”, In Proceedings of the ICPhS conference, 2437–2440, 2003.

[8] Droua-Hamdani, G., Selouani, S.A., Boudraa, M. and Cichocki,W., “Algerian Arabic rhythm classification”, ISCA International Speech Communication Association, in Proceedings of the third ISCA Tutorial and Research Workshop Experimental Linguistics, ExLing 2010, 25 - 27 August 2010, Athens, Greece, pp. 37-41, 2010.

[9] Selouani, S-A., Alotaibi, Y.A. and Pan, L., “Comparing Arabic rhythm metrics among other languages”, In: Audio, Language and Image Processing (ICALIP)”, 2012 International Conference, pp. 287-291, 2012.

[10] Droua-Hamdani, G., Alotaibi, Y.A., Selouani, S-A. and Boudraa, M., “Rhythmic Features across Modern Standard Arabic and Arabic Dialects”, Workshop on Free/Open-Source Arabic Corpora and Corpora Processing Tools. pp 43-47. 27 May 2014. (Reykjavik). Island. 2014.

[11] Droua-Hamdani, G., Selouani, S-A. and Alotaibi, Y.A., “Rhythm analysis in Arabic L2 speech”, Speech Prosody 07. pp.1007-1011. May 20--23, 2014 (Dublin). Ireland. 2014.

[12] Ramus, F., “Acoustic correlates of linguistic rhythm: Perspectives”. In Speech prosody, 115–120. Citeseer, 2002.

[13] Dellwo. V., “Rhythm and speech rate: A variation coefficient for deltaC”, Language and language processing, 231–241, 2006.

[14] Grabe, E. and Low, E.L., “Durational variability in speech and the rhythm class hypothesis”, Papers in laboratory phonology, 7:515–546, 2002.

[15] Bertinetto, P.M. and Bertini, C., “On modelling the rhythm of natural languages”, Proc. 4th International Conference on Speech Prosody, 427-430, Campinas, 2008.

[16] Droua-Hamdani, G., Selouani, S-A. and Boudraa, M., “Algerian Arabic Speech Database (ALGASD): corpus design and automatic speech recognition application”, The Arabian Journal for Science and Engineering 35, Number 2C, 157-166, 2010.

[17] Linguistic Data Consortium LDC. http://www.ldc.upenn.edu.