An acoustic analysis of the rhythm of Yemeni Arabic

Nada Mohammed Salem

nadasalem@siswa.um.edu.my

University of Malaya, Malaysia

Stefanie Pillai

stefanie@um.edu.my

University of Malaya, Malaysia

Abstract

Previous studies have found that different Arabic dialects display different degrees of stress-timing features forming a continuum that ranges from more stress-timed to less stress-timed Arabic dialects. This study investigated the vocalic durational intervals of speakers from two groups of speakers, Yemeni younger generation (YYG) and Yemeni older generation (YOG), both in reading and spontaneous speech by means of two indexes, namely, *n*PVI and Varco V. The results revealed that Yemeni speakers displayed stress-timing features in both the read passage and spontaneous speech contexts. However, whereas the rhythmic patterns in their read speech appeared to be comparable for both YYG and YOG speakers, significant differences were found in their spontaneous speech. This was in terms of the *n*PVI and VarcoV measurements where the rhythmic output of YYG speakers appeared to be more stress-timed than their older counterparts. The high values obtained by both indexes in the study show that YA patterns with prototypical stress-timed languages such as English, Dutch, Thai and the highly stress-timed western Arabic dialects.

Keywords: rhythm, stress-timed, syllable-timed, Yemeni Arabic, Arabic dialect

1.   Introduction

Standard Arabic (henceforth SA) is the official language of Yemen, but Yemeni Arabic (henceforth YA) is the language spoken by all Yemenis except for citizens in the governorates of Mahra and Socotra, who speak different languages that are unintelligible to the speakers of YA. SA is not regarded as the mother tongue of Yemenis. Rather, they are brought up using their regional variety, namely, YA, in all formal and informal occasions. Prior to 1962, Northern Yemen was under the imamate rule for over a thousand years. North Yemen was kept in deliberate insulation (Burrowes, 2010; McLaughlin, 2007; Rabi, 2014) until a republican rule was established (King, 2012). The establishment of new educational institutions in the 1970s and 1980s was accompanied by the need for experts in the field of writing curricula and teaching different subjects who came from different parts of the world. The increasing contact and interactions that Yemenis had with non-Yemenis, particularly Arabic speakers from different countries, may have influenced speakers' linguistic practices. Thus, there may be variation in the linguistic patterns of older and younger Yemenis. This study looks at the rhythmic patterns of this dialect as produced by two age groups. The questions that this study addresses are:

1.   What are the patterns of rhythm in Yemeni Arabic?
2.   To what extent is there a difference between the rhythmic patterns of the younger generation of Yemenis compared to the older generation?

2.   Speech Rhythm

Rhythm is one of the components of the prosody of human language which is described by Handel (1989) as an experience that "involves movement, regularity, grouping, and yet accentuation and differentiation" (p. 383). It is a prosodic feature that results from the repeated intervals that occur successively in human speech and is dependent on certain factors such as the syllable structure of the language, contrasts of short and long vowels, vowel reduction or the occurrence or absence of sequences of vowels (Ladefoged, 2006).

Abercrombie (1967) defined rhythm as "periodic recurrence of some sort of movement, producing an expectation that the regularity of succession will continue" (p. 96). This movement mainly involves stress and syllables and the recurrence of these two elements of speech results in what has been described by Pike (1945) as stress-timed or syllable-timed rhythm. This implies that rhythm is a prosodic feature of human speech that results from isochrony, namely, the regular

occurrence of units of speech like syllables or stress. Thus, rhythm is sensitive to the successive nature of features like syllables and stress, which helps to classify languages into stress-timed and syllable-timed languages.

Stress-timed rhythm displays considerable variations in terms of syllabic durations and structures and can be recognized in languages like English, Russian, Arabic and German. In other words, in these languages, stressed syllables recur at equal intervals, that is, they tend to be synchronous. The syllable-timed rhythm, on the other hand, shows less variation in terms of syllabic duration and more regularity in syllabic structures. That is to say, the syllables occur at equal intervals or are isochronous, as in languages like French, Spanish and Telugu (Abercrombie, 1967).

Later researchers argued against the dichotomic nature of rhythm and considered rhythm to be determined by the absence or presence of certain phonological and phonetic features of a language (e.g., Dasher & Bolinger, 1982; Dauer, 1983; Nespor, 1990). In studying the inter-stress intervals in English, Thai, Spanish, Italian and Greek, Dauer (1983) found that stress recurs more or less regularly in a prototypical stressed-timed language like English and a prototypical syllable-timed language like Spanish. Despite the more or less regular recurrence of stress in both English and Spanish, their rhythm appeared to be different. Therefore, Dauer (1983) suggested that rhythm output is an automatic consequence of certain linguistic circumstances that occur due to the interaction of a number of factors in a language, namely, the syllable structure, vowel reduction and stress. Stress-timed languages have a complex structure and frequent cases of vowel reduction, while syllable-timed languages display more constraints, and mora-timed languages showing even much more constraints. In the prototypical stress-timed languages like English, Swedish and Russian, vowels tend to be reduced in unstressed syllables resulting in the irregularity of the syllable structure of these languages. A fundamental factor in making the difference between durations of stressed and unstressed syllables in stress-timed languages greater and more evident is the extent to which vowels are reduced. Stress-timed languages also display lexical stress that results in an evident lengthening of the syllable that receives the stress, and that syllable is regarded as the turning point for intonation. The nucleus in those languages can consist of short and long vowels as well as diphthongs, and the syllable coda can bear up to four consonant clusters. On the contrary, in the prototypical syllable-timed languages there is no vowel reduction and they do not display a variety of syllable types. An example of this is the Spanish syllable structures that consist

only of a vowel or diphthong and the syllable can only allow for two consonants word medially (Dauer, 1983). Therefore, languages should not be treated as either stress-timed or syllable-timed. Rather, they should be thought of as forming a continuum in which there might be languages that occur in the polar opposites of the continuum with obvious stress-timing or syllable-timing features, while others may stand mid-way between the two extremes for displaying mixed features of rhythm.

Dauer's (1983) study of rhythm showed that certain phonological features such as syllable structure and phonotactic structure have an influence on the rhythmic structure of that language. Dauer's ideas were quantified by proposing indexes that enabled researchers to apply phonetic measurements that may provide phonetic validation of that phonologically based view of rhythm. Thus, after the prevalent perceptual analysis of rhythm in the past, a breakthrough came, in the late 1990s, when rhythmic indexes were used for instrumental analysis that supported the notion of rhythm. In other words, new and advanced instrumentations for measuring rhythm were proposed with the advent of rhythmic indexes. These include measuring the proportion of vocalic intervals (%V), standard deviations of vocalic intervals (ΔV) and consonantal intervals (ΔC) proposed by by Ramus, Nespor and Mehler (1999), the Pairwise Variability Index (PVI) proposed by Low, Grabe and Nolan (2000), the Variability Index (VI) proposed by Deterding (2001), the normalized version of the PVI (nPVI) proposed by Grabe and Low (2002) as well as VarcoC and VarcoV proposed by Dellwo (2006). Ramus, Nespor and Mehler (1999) applied their indexes to measure the vocalic and consonantal durations of English, French, Spanish, Italian, Polish, Dutch and Japanese. They found that Dutch and English appear to be grouped together as stress-timed languages for displaying high values on the scale of ΔC and ΔV and low values on the scale of %V. Catalan, French, Italian and Spanish, on the other hand are grouped together as syllable-timed languages with low values on the scale of ΔC and ΔV and high values on the scale of %V. As for Japanese, it appears in a separate and distinct position on all scales. Polish displays similarity with the English and Dutch values on the scales of ΔC and %V. However, such similarities are not seen between their values on the scale of ΔV.

The PVI index was applied to compare Singapore English with British English. Based on the results of the PVI values, Singapore English appeared to be more syllable-timed than British English due to the slightly longer or less reduced vowels found in the former. Grabe and Low (2002) applied the PVI, but with a new dimension to the index. The authors used the normalized

index or *n*PVI to measure the durational intervals of vowels that seemed to be sensitive to the rate of speech. This is unlike the durational intervals of consonants that is not sensitive to the rate of speech and can, therefore, be measured using the raw PVI. Using the *n*PVI, Grabe and Low (2002) calculated the durational intervals of the vowels of eighteen languages providing a continuum of languages that stretched from Thai with typical stress-timing features to Chinese with typical syllable-timing features. However, some languages stood midway in the continuum as they displayed mixed features of rhythm suggesting that rhythm is not dichotomous and is thus, better viewed in terms of a continuum that ranges from more to less stress-timing or syllable-timing features.

Later studies on rhythm modified those indexes to introduce other indexes such as 'Rhythm Ratio' (RR) which is a syllabic PVI proposed by Gibbon and Gut (2001). They compared the durations of successive syllables of speakers of British English, Nigerian English and Ibibio which is a syllable-timed Nigerian language. They found that Ibibio English is more syllable-timed than English, whereas Nigerian English showed more durational variations than British English. However, these findings are only based on the speech of one British speaker. Gut (2005) applied the RR index to investigate the rhythm of British English, Nigerian English and three Nigerian Languages (Yoruba, Hausa and Igbo). British English displayed the lowest RR value, followed by Nigerian English, whereas the Nigerian languages scored high RR values. This indicates that British English displayed the greatest variations in the durations of successive syllables and Nigerian English displayed less variations while the Nigerian languages displayed the least durational variations.

Prior to Gibbon and Gut (2001) also measured the syllabic duration pairwise using the Variability Index (VI) to compare the rhythm of British English and Singapore English. However, unlike the PVI and RR where the normalization of the index is achieved locally (interval variations that follow each other), the normalization in VI is global since the duration of each syllable is divided by the mean duration of all syllables and not only the pair of syllables as in PVI and RR. The findings revealed that Singapore English displayed significantly less variations in the durations of syllables and showed more regularity than British English which indicates that both varieties are not similar in the stress-timed/syllable-timed classification.

The syllabic duration was also the main focus in the rhythmic analysis of German, English, French and Italian by Wagner and Dellwo (2004) who proposed another normalization of the

syllabic PVI and called it YARD (Yet Another Rhythm Determination). Their results showed that English and German syllables displayed more variations with less evidence of isochrony that ranged from two to three syllables compared to French and Italian syllables which displayed less variations with more evidence of isochrony that ranged from three to four syllables.

All these studies lean towards the assumption that it is possible to discriminate between languages or dialects based on their rhythmic structures. Thus, several languages have been investigated and discriminated using these indexes or even by a combination of two or more indexes. In short, the rhythmic indexes are different means of describing and analysing rhythm acoustically by quantifying the differences of the rhythm of languages that result from the phonological, phonetic, syntactic and lexical features of those languages. It is worth mentioning that although these indexes remained very popular in the studies of speech rhythm, they may not be very reliable in classifying languages due to factors such as the involvement of speech rate that cannot be isolated in the acoustic analysis of rhythm. Russo and Barry (2008) pointed out that in applying the indexes in different speaking rates, low scores result from fast speech as in some Italian varieties. Thus, many languages that are classified as syllable-timed are spoken with faster speaking rates than the prototypically classified stressed-timed languages. Another major factor that affects the studies of rhythm is the direct impact of eliciting methods and inter-speaker variation that may result in different values for different speech samples. These differences can be are so prominent that they can result in different classifications for the same languages as shown in Arvaniti (2012. Such inconsistencies are quite problematic but unsurprising given the various factors that affect the durational intervals in speech. An unexpected finding was the more robust and consistent scores in the indexes that do not normalize for speech rate, such as %V, ΔC and rPVI. This is in comparison to Varcos and nPVI indexes that do normalize for speech rate. Generally, the classification of languages based on rhythmic indexes continues to be risky and, in many cases misleading (Arvaniti, 2012). However, this problem seems to be more serious in non-prototypical languages or languages that have not yet been classified, for instance Korean, probably due to the change in its prosodic system across generations. It is therefore important to examine sufficient speakers of a language in different speaking styles to try to overcome the problems that may result from methodological decisions, and thus, obtain to some extent, robust and reliable metric scores. It has to be noted though, that speakers in Arvaniti (2012) were asked to read utterances that were intentionally designed for different rhythms in English, German,

Spanish, Italian, Greek and Korean. The manipulated stress-timed Spanish utterances for instance gave scores comparable to those of the uncontrolled English sentences on the scale of ΔC and %V. It can be regarded as a serious issue if we look at rhythm as a property of language and not a property of speech. Therefore, although the materials contained structurally correct sentences from the investigated languages, they were designed for particular features of rhythm that are unlikely to occur in casual speech where no prevalence of a deliberate and exclusive type of rhythm is used (Fuchs, 2016).

2.1. Arabic Rhythm

In Arabic studies, earlier analysis of rhythm was also based on auditory impressions describing different Arabic dialects as being stress-timed as suggested by Hassan (1955). With the advent of rhythmic indexes, researchers like Barkat (2000), Hamdi-Sultan, Barkat and Ferragne (2004) and Ghazali, Hamdi-Sultan and Knis (2007) began to apply these indexes in acoustic analyses of Arabic rhythm. These studies conform to earlier perceptual studies regarding the classification of Arabic with stress-timed languages. However, these studies concluded that different degrees of stress-timing features are displayed by the different Arabic dialects. The western dialects like Moroccan and Algerian are considered to be more stress-timed than the eastern dialects such as Jordanian and Syrian.

In trying to situate SA in the Arabic rhythm continuum, Droua-Hamdani, Cichocki, Selouani and Boudraa (2010) investigated the speech timing of SA spoken by Algerian speakers. Interestingly, although the Algerian dialect belongs to the western group of Arabic dialects with high stress-timing features, in reading SA, the Algerian speakers displayed mixed features where their SA reflected some stress-timing features. However, the location of SA was closer to prototypical syllable-timed languages like French.

In another study, Hamdi-Sultan, Barkat, Ferragne and Pellegrino (2004) examined the rhythm of different Arabic dialect to see whether these dialects displayed similar rhythmic output, or if they ranged along a more stress-timed to less stress-timed continuum. It was revealed that in the dialects of the west, like Algerian and Moroccan, the vowels are frequently reduced and display more durational variations and complex syllable structures, like English which is typically classified as a stress-timed language as seen in Figure 1. On the other hand, although they are considered as stress-timed, the dialects in the east, like Syrian, Jordanian and Egyptian, do not

show frequent reductions and durational variations of vowels and syllable complexities. As for Tunisian, it falls midway between the dialects of the east and the west displaying mixed features of both groups of dialects. The vowels in Tunisian are longer than those of the western dialects but not as long as eastern dialects vowels, and the syllable structure is less complex than Moroccan and Algerian syllables. Variations in rhythm were found in Hamdi-Sultan *et al*. (2004) to be directly proportional to the distance between the areas in which languages or language varieties are used.
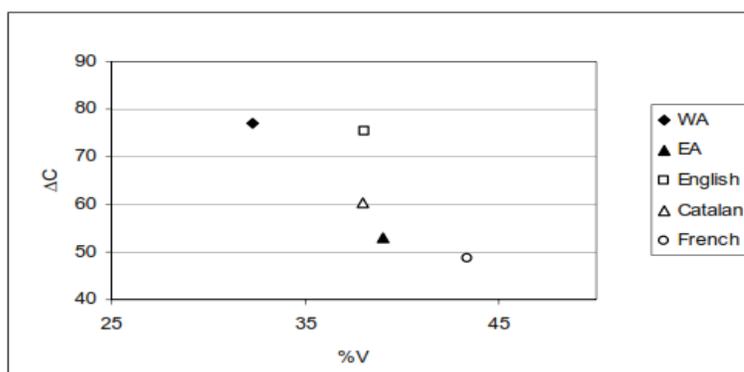

Figure 1: Distribution of Arabic dialects compared to English, Catalan and French on %V and ΔC scales
(Hamdi-Sultan *et al*., p. 1615)

In the same vein, Ghazali, Hamdi-Sultan and Knis (2007) found that the western dialects of Arabic are indeed more stress-timed than the eastern dialects. The %V scores showed a gradual rise from the western dialects to the eastern dialects while the scores of ΔC consistently decreased which indicates that the different groups of dialects show different degrees of stress-timing features. A comparison of the eastern dialects showed no significant differences among the values of those dialects indicating that this group of dialects displays comparable rhythmic patterns. However, significant differences were seen in comparing the eastern group with the western group of dialects as shown in Figure 2.

It has to be noted that the Egyptian and Tunisian dialects scored intermediate values between the eastern and western dialects although showing more inclination towards the rhythmic features of eastern dialects particularly in terms of ΔC values which showed significant differences when compared to western dialects. Similar results were found in Barkat (2000) where the Tunisian dialect displayed similarity to eastern dialects in terms of ΔC due to reduced complexity in the syllable structures compared to the western dialects. However, in terms of %V, the Tunisian

dialect showed similarity to western dialects. Although Tunisian vowels are longer than the latter, they show more reductions than the eastern dialects which make them comparable to the western dialects.
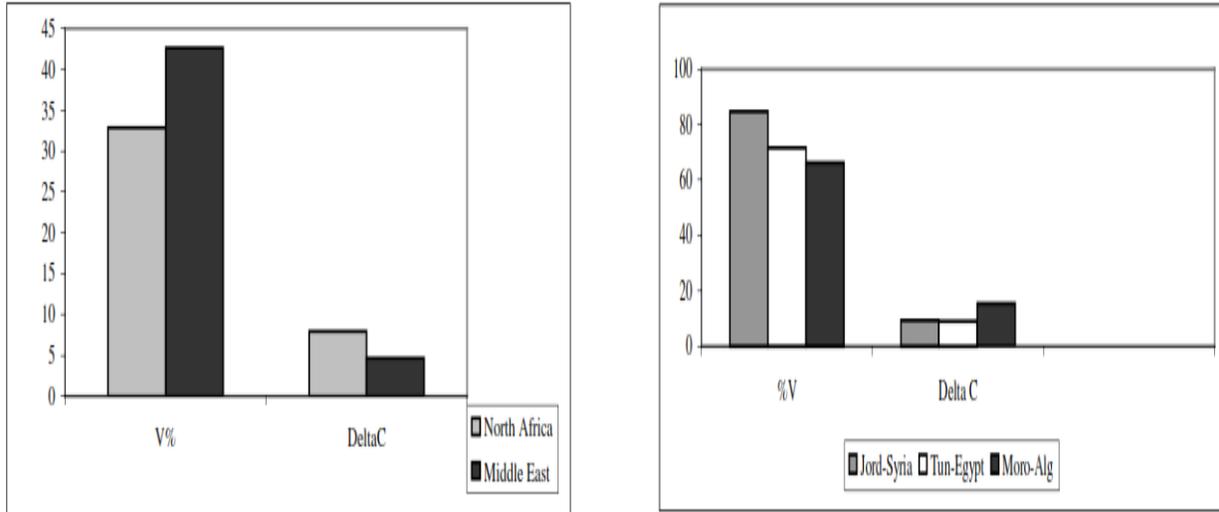
Figure 2: Comparison between Arabic dialects in (Ghazali *et al*., 2007, p. 333)

All these studies have attempted to place different Arabic dialects in the rhythm continuum. Thus, by investigating the patterns of rhythm in YA, this study aims to identify the rhythmic features of this dialect and determine whether it belongs to the more stress-timed or the less stress-timed Arabic dialects.

3.   Method

3.1.   Participants

All the participants in the present study lived in the city of Taiz. Their parents and spouses were also from the city of Taiz. All the speakers in this study were born and educated in the city of Taiz in Yemen. They spoke only YA in their homes (a dialect that geographically belongs to the eastern dialects). However, speakers of the younger generation started studying SA in schools at all stages from primary to high school and English as a foreign language from the age of 12 to 13 years old. As previously mentioned, this study attempts to compare the prosody of two different generations of YA. The first group represented the present generation and comprised 15 undergraduate students from a local university. All the participants were females and they were between 20 to 25

years old. The second group that represented the older generation comprised another 15 Yemeni speakers aged between 60 to 65 years. They had always lived in the city of Taiz and received the kind of education which was available during the Imamate rule or before the 1962 revolution which was mainly based on religious schools in which children memorized the holy Quran. The social and political instability and the occurrence of a civil war in the country during the time in which the data were collected was the main reason of limiting the participants to female speakers as only females were available in the safe places during the data collection period. Comparing two different generations was deemed necessary due to the completely different social and educational situations of the two groups.

3.2.   Materials

Different speaking contexts were used to elicit data for this study. A read text was used to obtain the same utterances from all speakers. Spontaneous speech was elicited to examine if the same patterns of rhythm occurred in this more naturalistic context.

The speakers were audio recorded individually using a digital recorder (Zoom H2n) in a quiet curtained and carpeted room away from any noise or distraction. All the participants were aware of the recordings and the purpose of their participation. Prior to the recordings, the speakers were asked to provide written consent to be recorded and they were informed about their right to withdraw from the study if they wanted to. The participants requested that the recordings not be played back to other people except for the purpose of analysis for the study.

3.2.1. Read Passage

A read passage is an important requirement to get the participants to read the same utterances that can provide a consistent and comparable material. However, given the fact that this study focuses on a regional variety and avoids SA in which texts are usually written, Lim's (1996) method of creating a read passage from spontaneous conversation was replicated. In other words, a Yemeni speaker was asked to speak spontaneously about the weather in Yemen, and her monologue was orthographically transcribed and used as a read passage, (see Appendix A). This made the material more natural and spontaneous than a read passage and the speakers of both groups did not face difficulties in reading the passage although it has to be said that whilst there is no formal written form of YA, the speakers use SA script to write in YA.

Eight utterances were obtained from every speaker. A total of 96 tokens were measured for every speaker. In total, 1440 tokens of vowel durations were obtained from each group, resulting in 2880 vocalic durations measured for all speakers from both groups.

3.2.2. Spontaneous Speech

The speakers were asked to talk about one of the topics presented to them like a special event in their lives, ways of spending their spare time, a memorable vacation, work, cooking, marriage or studying. Given that people's interests are different, providing a number of different topics is important to give the participants the chance to select a topic that matches their interests, (Milroy & Gordon, 2003). The topics presented in this study were selected to match the general interests of female speakers. Although studying and working might not be a preferred topic for the older participants, the other topics seemed to be appropriate for women of all ages. The speakers were recorded when they felt prepared and comfortable, and after they had the time to think about what they wanted to say.

Ten utterances were extracted for each YYG and YOG speaker. A total of 150 utterances were analysed for each group resulting in an analysis of a total of 300 utterances for both groups of speakers.

3.3   Data Analysis

The rhythm of YA was analysed in this study by measuring the vocalic durations of the dialect using Praat (Boersma & Weenink, 2012). Measuring the vocalic durations provides better results and is more effective than the consonantal durations in studies of rhythm (Grabe & Low, 2002; White & Mattys, 2007; Wiget *et al*., 2010). The *n*PVI which is the normalized version of the PVI, is one of the most successful indexes of rhythm and has been found in many studies such as Low, Grabe and Nolan (2000) and Loukina, Kochanski, Shih, Keane and Watson (2009) to be quite robust and reliable in capturing the differences between languages or even between language varieties. However, Loukina, Kochanski, Shih, Keane and Watson (2009) found that a combination of two rhythmic indexes provides more effective results and more successful classification of languages. White and Mattys (2007) applied different indexes to compare between the rhythm of Dutch, English, Spanish and French. They found both *n*PVI and VarcoV to be better indexes in achieving a more evident distinction between and within the different types of rhythm. Therefore,

this study investigated the rhythm of YA by combining two different rhythmic indexes, namely *n*PVI and VarcoV, for the purpose of providing a better insight into YA rhythm. The *n*PVI was used to measure the durational intervals of successive pairs of vowels and then, the mean of those variations was calculated as follows:

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

'*m*' = 'number of items in an utterance'

'*d*' = 'duration of the *k*th item'

<div align="right">(Grabe & Low, 2002, p. 517)</div>

For VarcoV, the ΔV is divided by the average value of the vocalic durational intervals and then multiplied by 100 for every YYG and YOG speaker's utterance:

$$\text{VarcoV} = \frac{\Delta V}{\overline{\mathcal{V}}} \times 100$$

ΔV = standard deviation of vocalic intervals

$\mathcal{V}$ = mean duration of vocalic intervals

<div align="right">(White & Mattys, 2007, p. 508)</div>

Whereas the *n*PVI measures the durational variations of successive adjacent pairs of vowels, VarcoV measures the global variations in utterances. The results obtained from both groups were compared for significant differences using *t*-tests (paired and independent *t*-tests for the read passage and an independent *t*-test for spontaneous speech with an alpha level 0.05).

4.   Findings

4.1. Read Passage

The values of the durational vocalic intervals of the YYG and YOG speakers are presented in this section. Eight utterances were extracted for analysis and 1440 tokens of vowel durations were measured for each YYG and YOG group of speakers, with a total of 2880 tokens for both groups. The mean *n*PVI values obtained from the YYG and YOG speakers' recordings are shown in Table 1.

Table 1: Mean *n*PVI values of vowel durations of the read passage of YYG and YOG

YYG *n*PVI	YOG *n*PVI
55.8	52.9

As previously mentioned, the study employs VarcoV as a second index for measuring and analying rhythm. Table 2 presents the results of the mean values of VarcoV obtained from the YYG and YOG speakers, while Figure 3 presents the *n*PVI and VarcoV values of both groups of speakers. The speakers seemed to produce comparable values of vocalic durations measured by both indexes. The *n*PVI values of the YYG and YOG speakers' vocalic durations were compared using an independent *t-test* (equal variance) and the results showed no significant differences between the two groups of speakers ($t(28) = 1.13$, $p = 0.27$). The effect size value ($r = 0.2$) suggested a low practical significance. Similarly, no significant differences were found between the VarcoV values of the YYG and YOG speakers ($t(28) = 1.60$, $p = 0.12$). The effect size value ($r = 0.3$) suggested a moderate practical significance.

Table 2: Mean VarcoV values of the vowel durations of the read passage of YYG and YOG

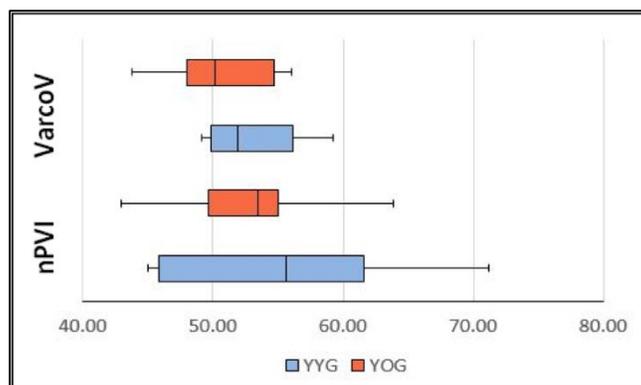| YYG VarcoV	YOG VarcoV
53.4	50.5

Figure 3: Boxplots of *n*PVI and VarcoV values of the vowel durations of the read passage of YYG and YOG

A paired *t-test* was also used to compare the values of the speakers since eight pairs of similar utterances were produced by the speakers of both groups. In other words, the mean *n*PVI values of the two groups for each of the utterances were compared and the results showed no significant differences, ($t(7) = 1.05$, $p = 0.31$). Similarly, no significant differences were found between the VarcoV mean values for each of the utterances ($t(7) = 1.25$, $p = 0.25$).

In the read passage, both YYG and YOG speakers produced comparable results in terms of PVI where high values (50 and above) were computed which denote that the language displayed more stress-timing features. Similar results were obtained from the VarcoV measurements which also did not display significant differences between the speakers of both groups. This means that that the variation of the durations of both YYG and YOG speakers were displayed in terms of successive pairs of vowels as captured by the PVI as well as the distribution of the vocalic durations in the entire utterance as captured by VarcoV.

4.2.   Spontaneous Speech

For spontaneous speech, the data were elicited from the speakers' conversation about one of the topics presented to them. 1,211 tokens were measured for YYG speakers and 1,107 tokens were measured for YOG speakers. A total of 2,318 tokens of vocalic durations were measured for YYG and YOG speakers. The *n*PVI values of the YYG and YOG speakers' vocalic durations are reported in Table 3.

Table 3: Mean *n*PVI values of the vowel durations of the spontaneous speech of YYG and YOG

YYG *n*PVI	YOG *n*PVI
60.15	57.48

Table 4 presents the results of the mean values of VarcoV obtained from the YYG and YOG speakers while Figure 4 presents the *n*PVI and VarcoV values obtained from the durational vocalic intervals of spontaneous speech utterances. The speakers seemed to display higher values than the read passage values indicating more stress-timing features in the spontaneous speech of both YYG and YOG speakers. An independent *t-test* was used to compare the *n*PVI values of YYG and YOG and the results showed significant differences between the two groups of speakers ($t(28) = 4.2$, $p = 0.0002$). The effect size value ($r = 0.6$) suggested a high practical significance.

14

The results suggest that the output of the YYG display more stress-timing features than their older counterparts showing more variations in terms of the duration of successive pairs of vowels.

Table 4: Mean VarcoV values of the vowel durations of the spontaneous speech of YYG and YOG
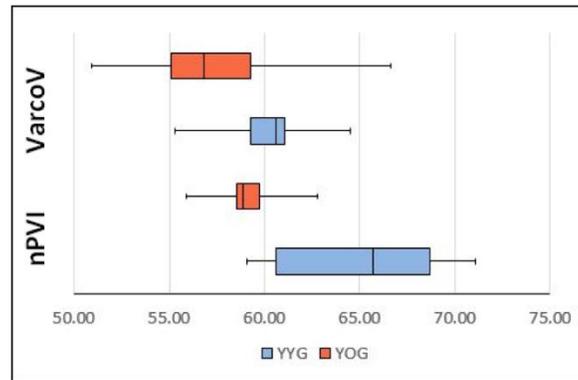
YYG VarcoV	YOG VarcoV
64.7	59.0

Figure 4: Boxplots of *n*PVI and VarcoV values of the vowel durations of spontaneous speech of YYG and YOG

Significant differences were found between the VarcoV values of the YYG and YOG speakers ($t(28) = 2.3$, $p = 0.03$). The effect size value ($r = 0.4$) suggested a moderate practical significance. Like the *n*PVI results, the VarcoV revealed that YYG speech output displayed more features of stress-timed rhythm than YOG speakers. Table 5 summarizes the results in the two speaking contexts.

Table 5: Summary of results

	Read Text		**Spontaneous Speech**	
GROUP	***n*PVI**	**VarcoV**	***n*PVI**	**VarcoV**
YYG	55.8	53.4	64.72	60.15
YOG	52.9	50.5	59.04	57.48

5. Discussion

We can infer from these findings that Yemeni speakers appear to display stress-timing features in both the read passage and spontaneous speech with all values higher than 50. This may be a reflection of certain phonological features related to vowel duration such as prominence or vowel reduction. As previously mentioned, several acoustic studies (see 2.1) have shown that Arabic dialect display increasingly more vowel reduction as one moves from the east to the west. The findings of this study, however, show that YA is considered as an exception for this rule. This is because high stress-timing features, comparable to the western dialects, characterize the speech of YA. Thus, despite belonging to the eastern dialects geographically, YA seems to belong to the western dialects rhythmically. However, it cannot be determined currently whether the high stress-timing features of YA, are caused by vowel reduction. This is because rhythm, according to Arvaniti (2012), cannot always be determined by means of a binary opposition in terms of the absence or presence of certain phonological features (Dauer, 1983). Therefore, it is possible that other processes may also be involved.

It is worth noting that the values of the spontaneous speech utterances seem to be higher, indicating that the speakers display more durational variation in those utterances, and thus, their stress-timing features were more manifested in spontaneous speech. The read passage utterances, although also displaying stress-timing features, showed less variability probably due to the influence of the reading task. In other words, although the text was created from spontaneous speech, the speakers were not as spontaneous and natural as they were in their spontaneous speech and their read speech produced fewer variations in their vocalic durations. These results are illustrated in Figure 5.

Figure 5 shows the distribution of YYG and YOG mean values computed by both *n*PVI and VarcoV indexes. A comparison of the distribution of the two groups of speakers in this figure to the distribution of languages in Grabe and Low (2002) suggests that YA rhythmic patterns are consistent with the prototypical stressed-timed languages. However, the spontaneous speech context appears to show more stress-timing features displaying high values comparable to typical stress-timed languages such as Thai, Dutch and German (Grabe and Low, 2002).
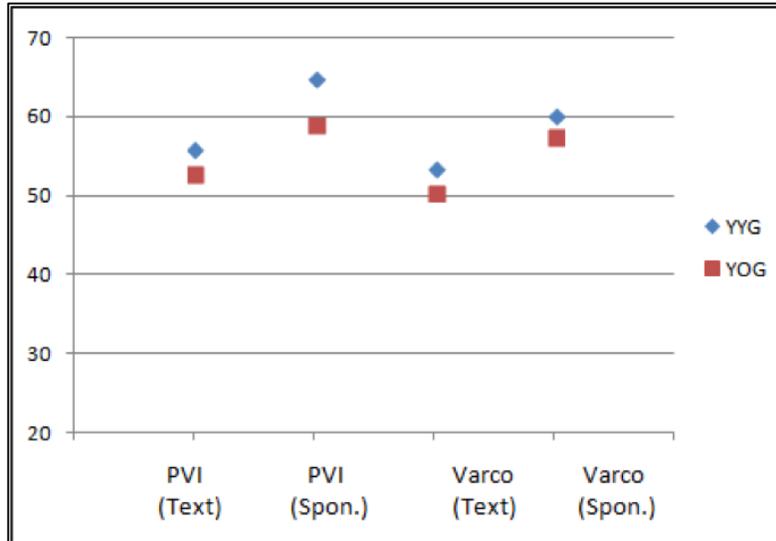
Figure 5: Distribution of YYG and YOG over the *n*PVI and VarcoV scales

Both the YYG and YOG groups produced comparable rhythmic patterns in producing the read passage. On the other hand, in the spontaneous speech, although both groups displayed stress-timing features, the YYG speakers' speech appeared to be more stress-timed than their older counterparts in terms of both PVI and VarcoV measurements where the results were significantly different between speakers of both groups.

In this study, both age groups experienced very different conditions and educational practices before and after the 1962 revolution. The different circumstances in which the speakers of both groups lived were expected to result in different realizations by both groups especially with the younger generation showing more inclination to adopt newer forms than the older generation (Trudgill,1988). Indeed, the speakers' spontaneous speech in this study shows that change over time is inevitable and it seems that, as Wolfram (2006) puts it, variation is the soul of language just like structure is its heart, for it is due to variations that nations, communities, social groups and individuals are distinguished. It is the heterogeneity and variation that ultimately bring about change in a language (Weinreich, Labov & Herzog, 1968). This change is experienced by a group of speakers who share the same stage in life and a place in history and is triggered by the experiences of individuals gained by their socialization and influenced by their race, gender, age, religion, ethnicity or social class (Eckert, 1997). However, the actual change in the rhythm of YA in this study occurs as a consequence of the change in time. In other words, age as a linguistic variable reflects change in the speech of Yemeni speakers as they move in time.

What seems to be interesting though, is the comparable results in the rhythmic features in the read passage. The YYG speakers are considered to be highly and differently educated compared to the older generation. Yet their reading patterns were surprisingly similar which is possibly due to the controlled content in the reading task. However, it may also indicate that the different educational practices did not have an influence on their reading skills. Thus, the differences may be due to their age difference rather than their different educational contexts. Perhaps a future study that examines the rhythmic patterns of younger and older generations in reading a SA text can confirm or contradict this notion.

6.   Conclusion

The results of this study revealed that Yemeni speakers appear to display stress-timing features in both the read passage and spontaneous speech. However, the rhythmic patterns in read speech are comparable for both YYG and YOG speakers. This is in contrast to their spontaneous speech, where there is evidence of stress-timing features by both YYG and YOG speakers. Significant differences were found between the results of both groups where the rhythmic output of YYG speakers appear to be more stress-timed than their older counterparts. Such differences are captured by both indexes denoting that variations are found in terms of successive vowel durations as well as the variant distribution of the vocalic durations in the entire utterance. However, these results may lead us to wonder why the significant differences were found in spontaneous speech and not in the read speech although the text was produced from a spontaneous speech of a speaker. It is possible that there is a real and actual difference between the younger and the older generations in their spontaneous speech, but it might also be an artefact of the different segmental content of the speech samples in spontaneous speech. In the read speech context, on the other hand, the segmental content was controlled, and therefore no age difference was found.

With reference to previous studies of Arabic rhythm (see 2.1), YA shows similar patterns with the western Arabic dialects due the high values obtained by both indexes applied in the study, and these values reflect the high stress-timing features of this dialect which are similar to those of the western dialects (e.g. Hamdi-Sultan *et al*., 2004, Ghazali *et al*., 2007 and Droua-Hamdani, Selouani, Boudraa & Cichocki, 2010). Moreover, although the two groups spoke the same dialect, they displayed significant differences in their rhythmic output which indicates that rhythm in YA

is undergoing change. Variations in rhythm were found in Hamdi-Sultan *et al*. (2004) to be directly proportional to the distance between the areas in which languages or language varieties are used. On the contrary, the two groups in this study displayed significant differences despite the lack of geographical divergence which Hamdi-Sultan *et al*. regarded as an important factor of variations in the rhythmic output of speakers.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351–373.

Barkat, M. (2000). Détermination des indices acoustiques robustes pour l'identification automatiques des parlers arabes (Unpublished doctoral dissertation). Université Lumière Lyon, France.

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.17, 2012. Retrieved 24 March 2012 from http://www.praat.org/.

Burrowes, R. D. (2010). *Historical dictionary of Yemen*. Lanham, MD: Scarecrow Press.

Dasher, R., & Bolinger, D. (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association*, *12*(2), 58-69.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*(1), 51-62.

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for DeltaC. In P. Karnowski & I. Szigeti (Eds.), *Language and language processing: Proceedings of the 38th linguistic colloquium* (pp. 231–241). Frankfurt: Peter Lang.

Deterding, D. (2001). The measurement of rhythm: A comparison of Singapore English and British English. *Journal of Phonetics*, *29*(2), 217–230.

Droua-Hamdani, G., Selouani, S. A., Boudraa, M., & Cichocki, W. (2010). Algerian Arabic rhythm classification. In Antonis, B. (Ed.), *Proceedings of the third ISCA tutorial and research workshop on experimental linguistics, ExLing 2010* (pp. 37-41). Athens, Greece.

Eckert, P. (1997). Age as a sociolinguistic variable. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 151-167). Oxford: Blackwell.

Fuchs, R. (2016). *Speech rhythm in varieties of English. Evidence from educated Indian English and British English.* Singapore: Springer.

Ghazali, S., Hamdi-Sultan, R., & Knis, K. (2007). Intonational and rhythmic patterns across the dialect continuum. In Benmamoun, Elabbas (Eds.), *Perspectives on Arabic linguistics* (pp. 97-122). Amsterdam: Benjamins.

Gibbon, D., & Gut, U. (2001). Measuring speech rhythm. In P. Dalsgaard (Ed.), *Proceedings of Eurospeech* (pp. 91–94). Aalborg, Denmark.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology Volume VII* (pp. 515-546). Berlin: Mouton de Gruyter.

Gut, U. (2005). Nigerian English prosody. *English World-Wide*, *26*(2), 153-177

Hamdi-Sultan, R., Barkat, M., Ferragne, E., & Pellegrino, F. (2004). Speech timing and rhythmic structure in Arabic dialects: A comparison of two approaches. *Proceedings of Interspeech 2004* (pp. 1613-1616). Jeju, Korea.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Hassan, T. (1955). *Linguistic methods in language research*. Cairo: Anglo-Egyptian Bookshop.

King, J. (2012). Zaydi revival in a hostile republic: Competing identities, loyalties and visions of state in Republican Yemen. *Arabica*, *59*(3-4), 404-445.

Ladefoged, P. (2006). *A course in phonetics* (5th ed). Los Angeles: Michael Rosenberg.

Lim, L. S. L. (1996). Prosodic patterns characterizing Chinese, Indian and Malay Singapore English (Unpublished doctoral dissertation). University of Reading, U.K.

Loukina, A., Kochanski, G., Shih, C., Keane, E., & Watson, I. (2009). Rhythm measures with language-independent segmentation. In *Proceedings of Interspeech 2009*. Brighton, U.K.

Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, *43*(4), 377-401.

McLaughlin, D. (2007). *Yemen*. Guilford, CT: The Globe Pequot Press.

Milroy, L., & Gordon, M. (2003). S*ociolinguistics: Method and interpretation*. Oxford: Blackwell.

Nespor, M. (1990). *On the rhythm parameter in phonology. The logical problem of language acquisition*. Dordrecht: Foris.

Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.

Rabi, U. (2014). *Yemen: Revolution, civil war and unification*. London: I. B. Tauris.

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265-292.

Russo, M., & Barry, W. J. (2008). Isochrony reconsidered. Objectifying relations between rhythm measures and speech tempo. *Proceedings of Speech Prosody 2008* (pp. 419-422). Campinas, Brazil.

Trudgill, P. (1988). Norwich revisited. Recent linguistic changes in an English urban dialect. *English World-Wide*, *9*(1), 33-49.

Wagner, P., & Dellwo, V. (2004). Introducing YARD (Yet Another Rhythm Determination) and reintroducing isochrony to rhythm research. *Proceedings of Speech Prosody 2004* (pp. 227-230). Nara, Japan.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds,), *Directions for historical linguistics* (pp. 95-189). Austin, TX: University of Texas Press.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*(4), 501-522.

Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*, *127*(3), 1559-1569.

Wolfram, W. (2006). Variation and language, an overview. In K. Brown (Ed.), *Encyclopaedia of language and linguistics volume 2* (pp. 333-340). Oxford: Elsevier.

Appendix

READ TEXT

بصراحة الجو بتعز ما فيش زيه. يعني بالبرد مثلا يكون الجو محتمل كثير مش زي صنعا البرد فيبه لا يطاق. وحتى بالصيف مع انه بالفترة الاخيرة   زاد الحمى مرّة بس برضو يظل الجو روعة زيادة لما يجي مطر. تلاقي الكل يخرج برع. محّد يشتي يجلس داخل البيت.

biṣarāħa al-gaw bi-taʿez māfīš zayu yaʿni bil-bard mathalan yikūn al-gaw muħtamal kathīr muš zay sanʿā al-bard fībeh lā yuṭāq w-ḥata bi-ṣef maʿ ino bil-fatra-l-axīra zād al-ħama mara bas barḍo yidal al-gaw rawʿa ziyāda lama yigi maṭar tilāqi-l-kol yixrog baraʿ maħad yišti yigles dāxel al-bet

Frankly the weather in Taiz is unmatched. I mean in winter for instance, the weather is quite tolerable, unlike Sana'a where the cold is unbearable. Even in summer, although lately it became much hotter, but still the weather is amazing especially when it rains. You find everyone goes outside. No one wants to stay at home.

About the Authors

Nada Mohammed Salem is a PhD student at the University of Malaya, Malaysia. She is a lecturer at Taiz University in Yemen but has been on study leave for her doctoral studies. Her main area of interest is Phonetics and Phonology, especially the prosodic features of varieties of Arabic. Nada received a Bachelor in English Literature from Taiz University. She obtained a Master's in Linguistics from the University of Malaya.
Email: nadasalem@siswa.um.edu.my

Stefanie Pillai[+] is a Professor at the Faculty of Languages & Linguistics, University of Malaya. Her areas of research interest include Acoustic Phonetics. She has published her work in journals including the *Journal of the International Phonetic Association*, *Lingua*, *Studies in Second Language Acquisition*, *World Englishes, Language Sciences* and *Language & Communication*.
Email: stefanie@um.edu.my
[+]Corresponding author