

**AN ACOUSTIC INVESTIGATION OF THE RHYTHM OF
YEMENI ARABIC AND JORDANIAN ARABIC**

NADA MOHAMMED ABDO SALEM

**DISSERTATION SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF LINGUISTICS**

**FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2014

**UNIVERSITI MALAYA
ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: **Nada Mohammed Abdo Salem** (I.C/Passport No: **04117917**)

Registration/Matric No: **TGC 110040**

Name of Degree: **Master of Linguistics**

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

An Acoustic Investigation of the Rhythm of Yemeni Arabic and Jordanian Arabic

Field of Study: **Phonetics**

I do solemnly and sincerely declare that:

- 1) I am the sole author/writer of this Work;
- 2) This Work is original;
- 3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- 4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- 5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- 6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date

Subscribed and solemnly declared before,

Witness’s Signature

Date

Name: DR. TAN SIEW KUANG, RACHEL

Designation: SUPERVISOR

ABSTRACT

In this study, the rhythm of Yemeni Arabic is measured and compared with Jordanian Arabic. The Arabic language has long been classified as being a stress-timed language. However, previous studies found that different Arabic dialects display different degrees of stress timing features forming a continuum that ranges from more stress-timed western Arabic dialects to less stress-timed eastern Arabic dialects.

Yemeni Arabic is one of the Arabic dialects that has received little attention in terms of rhythm. Therefore, this study attempts to present an account of Yemeni rhythm based on the acoustic measures of its consonantal and vocalic durational intervals and compare them to those of Jordanian, which is one of the established and widely studied Arabic dialects. Recordings of Standard Arabic read speech and spontaneous Arabic speech of 10 Yemeni and Jordanian speakers were measured and analyzed using the Pairwise Variability Index (PVI) that calculates the durational variability in successive intervals. Based on the ideas posited by Dauer (1983), the measurements reflect certain phonological features, such as the syllable structure and vowel duration, that have an influence over the rhythmic structure of a language. The findings revealed that both Yemeni and Jordanian Arabic do not pattern differently in terms of the consonantal and vocalic intervals, thus, they seem to have comparable rhythmic structures. Interestingly, the two parameters used in this study resulted in two different classifications of rhythm types. That is to say, in terms of the consonantal durations, the PVI values of both Yemeni and Jordanian showed less durational variability indicating syllable-timed rhythm, whereas the PVI values of the Yemeni and Jordanian vocalic durations showed greater durational variability indicating stress-timed rhythm. The rhythm output of the speakers could be influenced by the mixed phonological features that these dialects display.

ABSTRAK

Dalam kajian ini, pengukuran irama bahasa Arab dialek Yaman dibuat dan kemudian dibandingkan dengan bahasa Arab dialek Jordan. Bahasa Arab telah lama diklasifikasikan sebagai bahasa yang dipengaruhi oleh tekanan pada suku-kata perkataan. Walaubagaimanapun, kajian sebelum ini mendapati bahawa dialek-dialek Arab yang berlainan iaitu dialek Arab daerah barat dan dialek Arab daerah timur mempunyai tahap tekanan yang berbeza pada suku-kata perkataan.

Bahasa Arab dialek Yaman adalah salah satu daripada dialek-dialek Arab yang kurang diberi perhatian terutama sekali berkenaan irama. Oleh itu, kajian ini bertujuan untuk menunjukkan irama dialek Yemen berdasarkan ukuran akustik tempoh selang masa konsonan dan vokal. Seterusnya, kajian ini juga bertujuan untuk membandingkan ukuran akustik ini dengan bahasa Arab dialek Jordan, iaitu salah satu daripada dialek Arab yang terkenal dan yang telahpun dikaji secara meluas. 10 rakaman bacaan dan juga pertuturan secara spontan dari penutur-penutur dialek Yemen dan Jordan telah diuji dan dianalisis menggunakan Pairwise Variability Index (PVI). PVI adalah satu kaedah untuk mengira variasi di dalam tempoh selang masa yang berturutan. Berdasarkan idea-idea yang dikemukakan oleh Dauer (1983), ukuran tersebut mencerminkan ciri-ciri fonologi tertentu, seperti struktur suku-kata dan tempoh selang vokal yang mempunyai pengaruh ke atas struktur berirama bahasa. Dapatan kajian menunjukkan bahawa kedua-dua bahasa Arab dialek Yemen dan Jordan tidak berbeza dari segi selang konsonan dan vokal. Justeru itu, kedua-dua dialek seolah-olah mempunyai struktur berirama yang hampir sama. Yang menarik sekali, kedua-dua parameter yang digunakan di dalam kajian ini menghasilkan dua klasifikasi jenis irama yang berbeza

Ini bermakna, dari segi tempoh masa konsonan nilai PVI kedua-dua dialek Yaman dan Jordan. Yang rendah menunjukkan rima jenis tekanan-suku-kata, manakala nilai PVI tempoh masa vocal yang tinggi menunjukkan rima jenis tekanan-masa.

ACKNOWLEDGEMENT

I would like to express my deepest thankfulness to Allah the Almighty for showering me with his blessings that enabled me to complete this modest work. I am sincerely and deeply grateful for my supervisor, Dr. Tan Siew Kuang, Rachel, for providing me with such helpful advices and insightful remarks which have been a real asset and true inspiration for me in each and every single step of doing my research. Special thanks to her for her patience and invaluable feedback and for her inspirational guidance and never-ending encouragement that established me as a seeker for knowledge.

With heartfelt gratitude and appreciation I acknowledge my indebtedness to my parents who pray for my success every single day and for their efforts and unconditional support to help me in completing my study. I am also thankful to my brothers and sisters and all my family members for their care and encouragement.

I would also like to extend my deep thanks to my husband for being there for me all the time with his love, patience and support that made this study possible.

Finally, thanks are due to all my lecturers and staff in the Faculty of Languages and Linguistics and to the participants of the study whose patience and cooperation were of utmost importance in conducting this study.

TABLES OF CONTENTS

DECLARATION	ii
ABSTRACT	iii
ABSTRAK	iv
ACKNOWLEDGEMENT	v
TABLES OF CONTENTS	vi
LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1 - INTRODUCTION	
1.1 Introduction	1
1.2 Background of the Study	1
1.3 Background of Yemen and Jordan	3
1.3.1 Background of Yemen (Republic of Yemen)	3
1.3.2 Background of Jordan (Hashemite Kingdom of Jordan)	4
1.4 Statement of the Problem	5
1.5 Purpose of the Study	7
1.6 Objectives of the Study	8
1.7 Research Questions	8
1.8 Scope and Limitations	9
1.9 Significance of the Study	9

1.10	Definition of Terms	11
1.10.1	Segmentals and Suprasegmentals	11
1.10.2	Rhythm	12
1.10.3	Stress-Timed Rhythm	12
1.10.4	Syllable-Timed Rhythm	12
1.10.5	Rhythmic Indexes	13
1.11	Outline of the Chapters	13
1.12	Conclusion	14

CHAPTER 2 – LITERATURE REVIEW

2.1	Introduction	16
2.2	Arabic Language and its Dialects	16
2.2.1	Syllable Structure of SA, YAD and JAD	20
2.2.2	Vowels of SA, YAD and JAD	22
2.2.3	Consonants of SA, YAD and JAD	24
2.3	Prosody	26
2.4	Speech Rhythm	30
2.4.1	Different Approaches to Rhythm	33
2.4.2	Dauer's (1983) Approach as a Theoretical Framework of the Study	39
2.4.2.1	Syllable Structure	42
2.4.2.2	Vowel Reduction	43
2.4.2.3	Stress	44
2.5	Rhythmic Indexes	46
2.6	PVI as an Analytical Framework of the Study	62

2.7	Speech Rhythm in Arabic	65
2.8	Conclusion	70
 CHAPTER 3 – METHODOLOGY		
3.1	Introduction	72
3.2	Research Design	72
3.3	Materials	74
3.4	Participants	75
3.5	Data Collection	77
	3.5.1 Passage Reading	78
	3.5.2 Spontaneous Speech	79
3.6	Data Analysis	81
	3.6.1 Analysis of Vowels	84
	3.6.2 Analysis of Consonants	87
	3.6.2.1 Stops	87
	3.6.2.2 Fricatives and Affricates	89
	3.6.2.3 Glide and Consonants	91
	3.6.2.4 Liquid Consonant /l/ and Trill /r/	92
	3.6.2.5 Nasal Consonants	92
	3.6.2.6 Pharyngealized Consonants (Emphatics)	94
3.7	Conclusion	95

CHAPTER 4 – RESULTS AND DISCUSSION

4.1	Introduction	96
4.2	Reading Text (SA)	96
4.2.1	Consonantal Durations of Yemeni and Jordanian Speakers	96
4.2.2	Vocalic Durations of Yemeni and Jordanian Speakers	98
4.3	Spontaneous Speech	101
4.3.1	Consonantal Durations of YAD and JAD speakers	101
4.3.2	Vocalic Durations of YAD and JAD speakers	103
4.4	Discussion	106
4.5	Conclusion	113

CHAPTER 5 – CONCLUSION

5.1	Introduction	114
5.2	Summary of Results	115
5.2.1	The Acoustic Correlates of the Rhythm of SA as Read by Yemenis and Jordanians	116
5.2.2	The Acoustic Correlates of the Rhythm of YAD and JAD in Spontaneous Speech	119
5.2.3	YAD Rhythm compared to JAD	121
5.3	Implications of the Study	124
5.4	Limitations of the Study	125
5.5	Recommendations for Further Research	126

REFERENCES	127
APPENDIX A: Read Passage Material	137
APPENDIX B: Translation of the Read Passage Material	138
APPENDIX C: Background of Yemeni Participants	139
APPENDIX D: Background of Jordanian Participants	140
APPENDIX E: Consent of Participants	141

LIST OF TABLES

Table 2.1: The SA Consonants (Adapted from Thelwall and Sa'addedin (2003, p. 51)	25
Table 2.2: Summary of the phonological features following Dauer's criteria	44
Table 4.1: Computed <i>r</i> PVI values of the consonantal durations of the SA read text of Yemeni speakers	97
Table 4.2 : Computed <i>r</i> PVI values of the consonantal durations of the SA read text of Jordanian speakers	98
Table 4.3: Computed <i>n</i> PVI values of the vocalic durations of the SA read text of Yemeni speakers	99
Table 4.4: Computed <i>n</i> PVI values of the vocalic durations of the SA read text of Jordanian speakers	99
Table 4.5: Computed <i>r</i> PVI values of the consonantal durations of YAD spontaneous speech	102
Table 4.6: Computed <i>r</i> PVI values of the consonantal durations of JAD spontaneous speech	103
Table 4.7: Computed <i>n</i> PVI values of the vocalic durations of YAD spontaneous speech	104
Table 4.8: Computed <i>n</i> PVI values of the vocalic durations of JAD spontaneous speech	104
Table 5.1: Summery of the PVI values	115

LIST OF FIGURES

Figure 1.1: Arabic Dialect Continuum	6
Figure 2.1: The Cardinal Vowel Chart of SA (Adapted from Thelwall and Sa'addedin (2003, p. 52)	23
Figure 2.2: Distribution of other languages over the <i>n</i> PVI and <i>r</i> PVI scales from Grabe and Low (2002)	46
Figure 2.3: Distribution of languages over (Δ C), (Δ V) and V% scales (Ramus <i>et al.</i> , 1999, pp30 32)	49
Figure 2.4: Differences between the languages based on <i>n</i> PVI and <i>r</i> PVI values (Grabe and Low, 2000)	53
Figure 2.5: The PVI Values of Nine Languages, (Ladefoged, 2006, p. 246)	54
Figure 2.6: Discrimination between languages based on Varco C and Δ C scales (Dellwo, 2006, pp. 232-235)	56
Figure 2.7: Difference between the Standard Deviation and the PVI (Mairano 2010, p. 40)	63
Figure 2.8: Comparison of the eastern and western Arabic dialects with other Languages (Hamdi <i>et al.</i> , 2004)	68
Figure 2.9: Distribution of Arabic dialects on %V and Δ C scales (Ghazali <i>et al.</i> , 2002)	70
Figure 3.1: Spectrogram of the production of the sound /l/ by Yemeni speaker 4	82
Figure 3.2: Spectrogram of the production of the sound /u/ by Yemeni speaker 4	82

Figure 3.3: Spectrogram of the production of the sound /aa/ by Jordanian speaker 5	86
Figure 3.4: Spectrogram of the production of the sounds /k/ and /d/ by Jordanian speaker 1	88
Figure 3.5: Spectrogram of the production of the sound ‘ʔ’ by Jordanian speaker 2	89
Figure 3.6 : Spectrogram of the production of the sounds /x/ and /dʒ/ by Jordanian speaker 4	90
Figure 3.7: Spectrogram of the production of the sound /j/ by Yemeni speaker 2	91
Figure 3.8: Spectrogram of the production of the sounds /r/ and /l/ by Yemeni speaker 2	92
Figure 3.9: Spectrogram of the production of the sounds /n/ and /m/ by Yemeni speaker 1	93
Figure 3.10: Spectrogram of the production of the sound /s ^ʕ / by Jordanian speaker 1	95
Figure 4.1: Computed <i>r</i> PVI values of the consonantal durations of Yemeni and Jordanian Speakers of the read text	100
Figure 4.2: Computed <i>n</i> PVI values of the vocalic durations of Yemeni and Jordanian Speakers of the read text	100
Figure 4.3: Computed <i>r</i> PVI values of the consonantal durations of YAD and JAD Speakers in spontaneous speech	105
Figure 4.4: Computed <i>n</i> PVI values of the vocalic durations of YAD and JAD Speakers in spontaneous speech	105
Figure 4.5: Spectrogram of the vowel contrasts in Arabic	110
Figure 4.6: Distribution of YAD and JAD and other languages over the <i>n</i> PVI and <i>r</i> PVI scales	111

LIST OF ABBREVIATIONS

C	Consonant
IIUM	International Islamic University Malaysia
IPA	International Phonetic Association
IPS	Institute of Postgraduate Studies
JAD	Jordanian Arabic Dialect
<i>n</i> PVI	normalized Pairwise Variability Index
PVI	Parwise Variability Index
<i>r</i> PVI	raw Pairwise Variability Index
SA	Standard Arabic
UKM	University Kebangsaan Malaysia
UM	University of Malaya
V	Vowel
YAD	Yemeni Arabic Dialect
YARD	Yet Another Rhythm Determination

CHAPTER 1

INTRODUCTION

1.1 Introduction

The purpose of this study is to investigate the acoustic correlates of the rhythm of Yemeni Arabic dialect (henceforth YAD). Given that comparing between languages and dialects is a common practice in the field of phonetics, the acoustic correlates of YAD rhythm are compared to those of Jordanian Arabic dialect (henceforth JAD) which is one of the most established and widely studied dialects of Arabic. This chapter introduces the background of the study. It proceeds with the statement of the problem, purpose and objectives of the study, the research questions, the significance of the study and finally the scope and limitations of the work. The key terms which are used in the study are also defined. Note that all the phonetic symbols used in describing the language are adopted from the work of Thelwall and Sa'addedin (2003) of describing the Arabic sounds in the *Handbook of the IPA*.

1.2 Background of the Study

Rhythm is an essential part of our lives. However, there has been no single definition of the term since it is a multidisciplinary topic that is manifested in different forms and used in various fields such as music and poetry; it is even used in science for describing the regular occurrence of particular physiological functions in the body. The different kinds of rhythm found in different aspects of life share a measurable nature of the flow of movement or

beat. Apart from these types of rhythm, the term in this study refers to speech rhythm which is perceived as temporal intervals that occur successively in a speech context.

Rhythm has been increasingly studied over the past decades and researchers started to investigate the aspects of rhythm in speech yielding several approaches to rhythm that perceived it as a result of short and long syllables, accent and stress or timing. However it was Pike's (1945) remarks, about the stress-timed and syllable-timed rhythm, followed by those of Abercrombie (1967) that laid the foundation of a different approach proposing that languages can be discriminated based on the stress-timing or syllable-timing features of their rhythm. This was a motivation for many researchers to conduct later on further studies on rhythm of languages to investigate the extent to which this assumption of discrimination in terms of rhythm is accurate. However Dauer (1987) suggested that rhythm should not be treated as being dichotomous since there are languages that display mixed features of both stress and syllable-timed rhythm. As such, it might be more useful and fruitful to compare between languages or dialects instead of merely discriminating between them.

The description and analysis of rhythm in the past was mainly based on perceptual findings. However, in the late 1990s, a breakthrough came in when rhythmic indexes were used for instrumental analysis that supported the notion of rhythm. Some of these indexes are the (%V), (ΔV) and (ΔC) proposed by Ramus, Nespors and Mehler (1999) and the Pairwise Variability Index (PVI) proposed by Low, Grabe and Nolan (2000). The rhythmic indexes were different means of describing and analyzing rhythm acoustically by quantifying the differences of the rhythm of languages that result from the phonological, phonetic, syntactic and lexical features of those languages. By applying these indexes in investigating rhythm,

a number of languages have been discriminated and classified based on their rhythm such as English and Russian as stress-timed languages and Spanish and Italian as syllable-timed languages.

As for Arabic language, it has always been described as being stress-timed. However, these conclusions in the past were only based on perceptual studies and after the development of the rhythmic indexes, the acoustic studies came in line with the previous perceptual findings suggesting that Arabic is indeed a stressed-timed language. Yet, some disparities have been found among different Arabic dialects like Jordanian, Syrian, Lebanese, Egyptian, Tunisian, Moroccan and Algerian. Although these dialects appeared to be stress-timed, they displayed different degrees of stress-timing rhythm. As such, the Arabic dialects are treated in terms of a continuum that ranges from the east to the west with the dialects in the west displaying higher degrees of stress-timing features than those in the east.

1.3 Background of Yemen and Jordan

1.3.1 Yemen (Republic of Yemen)

Yemen is an Arab country that is located in the southern end of the Arabian Peninsula. It is the second largest country in the Peninsula after Saudi Arabia. The Roman called Yemen as Arabia Felix (Happy Arabia) for being blessed with wide spaces of greenery that distinguishes it the other countries in the Peninsula due to rainfall, fertile fields and lands that support the cultivation of different kinds of plants, (Resto, 2000, p. 189). Yemen has been divided into twenty two governorates since 2004, however, in 2014 it became a

federal republic with six regions. It is inhabited by about 24 million citizens¹. Whereas Standard Arabic (henceforth SA) is the official language, YAD is the language spoken in all Yemen except for Mahra² and Socotra³, where completely different languages are spoken there which are unintelligible to the speakers of YAD.

1.3.2 Jordan (Hashemite Kingdom of Jordan)

Jordan is an Arab kingdom known as the Hashemite Kingdom of Jordan. It is one of the Middle Eastern countries that is located in the south of the Levant region. Jordan is one of the largest producers of phosphate in the world. Besides this natural resource, it is known to be a destination of medical tourism in the Middle East.

The population of Jordan is about 6,249,000, however, not all the citizens are originally Jordanians. Many of them are of Palestinian origins that migrated to Jordan after the Arab-Israeli war in 1948. Today, there are about 1,951,603 Palestine refugees who have been offered Jordanian nationalities. Other immigrants include about 1,000,000 Iraqis, 15,000 Lebanese and 500,000 Syrian refugees. Like the other Arabic countries, SA is the official language in Jordan but the Jordanian dialect of Arabic is the language spoken in the country.

¹ <http://www.sabanews.net>

² Mahra is a Yemeni governorate that is located in the extreme east of Yemen

³ Socotra is a group of 4 islands that belong to Yemen and are located in the Indian Ocean

1.4 Statement of the Problem

Rhythm has been a widely investigated topic until today, yet, we do not have a thorough description and analysis of the rhythm of Arabic dialects. Studies on Arabic have always been focusing on dialectical variation in terms of segmental features. It was only in the late 1990s that researchers moved towards the suprasegmental variations across the different Arabic dialects. In studying the Arabic rhythm, researchers like Miller (1984), Tajima, Zawaydeh and Kitahara (1999) have categorized Arabic with stress-timed languages based on their impressionistic studies. However, after the advent of new rhythm indexes like those proposed by Ramus *et al.* (1999), Low *et al.* (2000) and Grabe and Low (2002), researchers moved towards a more successful acoustic investigation of rhythm. In their acoustic studies of Arabic rhythm, Barkat (2000) (as cited by Ghazali, Hamdi and Knis, 2007), Ghazali, Hamdi and Barkat (2002), Hamdi, Barkat, Ferragne and Pelegrino (2004) and Ghazali *et al.* (2007) all agree that Arabic is regarded as a stress-timed language. However, different Arabic dialects display different degrees of stress-timing features forming a continuum of dialects ranging from the east to the west with those in the latter displaying more stress-timed rhythm than the eastern dialects. The western Arabic dialects are represented by dialects of North Africa, namely, Morocco, Algeria and Tunisia, while Lebanese, Syrian, Jordanian and Egyptian dialects represent the eastern Arabic dialects as can be seen in figure 1.1:



Figure 1.1: Arabic Dialect Continuum (from Middle East Atlas)

According to some researchers like Ghazali *et al.* (2002), Hamdi *et al.* (2004) and Ghazali *et al.* (2007), there are some Arabic dialects that display features of both the eastern and western Arabic dialects. Hamdi *et al.* (2004) for instance, suggested that Tunisian is an intermediate dialect that displays vowels which are longer than those of Moroccan and Algerian dialects but shorter than the vowels of eastern dialects like the Syrian, Jordanian, Lebanese or Egyptian dialects. However, unlike Hamdi *et al.* (2004) who suggested that only Tunisian has mixed features from both sets of Arabic dialects, Ghazali *et al.* (2002), found that both Tunisian and Egyptian are regarded as intermediate Arabic dialects which display intermediate features of stress-timed rhythm. Thus, all these researchers dealt with this bipolar continuum as a description of the case of Arabic rhythm among the different Arabic dialects. However, various Arabic dialects like Yemeni, Saudi, Omani, Qatari, Bahraini, Emirati, Iraqi and Libyan have been left out; in other words, the rhythm of many Arabic dialects, to the best of the researcher's knowledge, has not yet been thoroughly investigated and classified with the eastern or western Arabic dialects. Thus, this study investigates the acoustic correlates of the rhythm of YAD and compares them to those of

JAD, one of the eastern dialects, as an attempt to locate the position of YAD among the previously investigated and classified Arabic dialects.

1.5 Purpose of the Study

The purpose of this work is to investigate, by means of spectrographic analysis, the acoustic correlates of the YAD Rhythm and to compare them to those of one of the eastern Arabic dialects in order to ascertain the position of YAD among the other classified Arabic dialects. Given that JAD is a dialect that has been the focus of much research, it can provide a basis for comparison, and thus, it was chosen in this study as a representative of the eastern Arabic dialects. YAD is one of the Arabic dialects which has not received much attention particularly in terms of rhythm. This paucity of research on the dialect in question drives this study. Thus, the work aims at exploring the aspects of rhythm in this dialect and identifies its position among the other investigated Arabic dialects.

In order to fulfill the required and intended aim, this study attempts to conduct measurements and analysis of recordings of Yemeni and Jordanian participants reading a SA text and then speaking in their dialectical or spontaneous speech. The SA text is considered to be useful for the study to measure similar and fixed utterances for all the participants providing a better scope for comparability.

The importance of investigating the spontaneous speech, on the other hand, lies in the fact that the ultimate goal of the study is to provide a description of an Arabic dialect, therefore, the analysis of spontaneous speech is inevitable in the study. Due to the comparative nature of phonetic research, a well-defined and researched dialect like Jordanian needed to be analyzed and compared with the dialect under investigation.

1.6 Objectives of the Study

The research objectives can be summarized as follows:

- (i) To investigate the acoustic correlates of the rhythm of SA as read by Yemenis and Jordanians.
- (ii) To investigate the acoustic correlates of the rhythm of spoken Yemeni Arabic dialect and Jordanian Arabic dialect.
- (iii) To compare the rhythm of Yemeni Arabic dialect with the rhythm of Jordanian Arabic dialect.

1.7 Research Questions

In order to fulfill the research objectives, this study aims to answer the following questions:

- (i) What are the acoustic correlates of the rhythm of SA as read by Yemenis and Jordanians?
- (ii) What are the acoustic correlates of the rhythm of spoken Yemeni Arabic dialect and Jordanian Arabic dialect?
- (iii) How does the rhythm of Yemeni Arabic dialect compare with the rhythm of Jordanian Arabic dialect?

1.8 Scope and Limitations

The study investigates the rhythmic structures of both YAD and JAD. Although there are a number of Arabic dialects that have been studied in terms of rhythm, JAD remains one of the most established and investigated dialects of Arabic. Therefore, it has been chosen as a representative of the eastern dialects and used for the sake of comparing its rhythm to that of YAD which is the subject of investigation in this work. Thus, only Yemeni and Jordanian speakers are targeted to provide the required data. Speakers of other Arabic dialects will not be included in the study.

This study investigates the rhythmic structure of a SA read text as well as spontaneous speech of Yemeni and Jordanian speakers to provide a clearer comparison between the rhythms of these two Arabic dialects. As such, the focus is mainly on speech rhythm, other segmental and suprasegmental features of speech are not investigated in the current study.

All the informants of this study are female postgraduate students in different universities in Malaysia. Although it was intended to collect data from a single educational setting, in practice, it was difficult to find a sufficient number of participants, who meet the requirements of the study from one university.

1.9 Significance of the Study

The main significance of the study is that it deals with one of the Arabic dialects that has received little attention particularly in terms of instrumental phonetics. Thus, it may

contribute to knowledge by its attempt to identify the rhythm of YAD as a first step towards exploring the rest of the dialects in the Arabian Peninsula, and thus, providing a more evident image of the Arabic dialects rhythm. Thus, the study is characterized by an exploratory nature and as Babbie (2005, p. 89) suggested, an exploratory study needs to fulfill the following requirements:

- (i) Provide the researcher with more knowledge and satisfies his/her inquisitiveness.
- (ii) Provide a test of feasibility for conducting a wider study.
- (iii) Provide appropriate methods that can be adopted in future work.

As such, the present study attempts to present a successful description and analysis of YAD and JAD to fulfill the above requirements and thus include YAD to the previously investigated Arabic dialects as well as pave the path for subsequent studies to explore the rhythmic structures of the dialects which are left uninvestigated yet. Moreover, unlike many previous studies which investigated Arabic rhythm by focusing on either a reading text or spontaneous speech, this study includes an investigation of both read SA and spontaneous speech of Yemenis and Jordanians providing a clearer comparison between the rhythm of these two Arabic dialects.

1.10 Definition of Terms

1.10.1 Segmentals and Suprasegmentals

In any speech context, two types of information are conveyed. The first is the segmental information which is related to those features of the language which are recognized as separated or ‘discrete’ segments, (Clark, Yallop and Fletcher, 2007, p. 326), in other words, the vowels and consonants of a particular language. The second information conveyed to a listener is that which is related to properties extending over the range of a single segment or the range of vowels and consonants. Laver (1994, p. 152) defined suprasegmentals as ‘factors which can potentially be prolonged beyond the domain of the segment’, such as pitch, rhythm, intonation, stress and tempo.

It is worth mentioning that some researchers like Ball and Rahilly (1999) and Crystal (2003) used the terms ‘suprasegmental features’ and ‘prosodic features’ synonymously referring to those features beyond the discrete vowels and consonants. Whereas other researchers like Fox (2000) and Trask (2007) viewed prosody to be different from suprasegmentals in the sense that the former is a phonetic concept and the latter is a phonological concept. Prosody is a term which is used specifically to refer to phonetic features which extend over the vowels and consonants while the suprasegmentals is regarded as a more general term that besides covering those features of prosody, it covers non-linguistic factors such as the voice quality that results from the nature of the larynx and vocal tract of the speaker, or paralinguistic factors such as nervousness or fatigue, Clark *et al.* (2007, p. 327). However, following Ball and Rahilly (1999) and Crystal (2003), the current study treats the two terms as being synonymous.

1.10.2 Rhythm

Rhythm is one of the prosodic or suprasegmental features of speech. Given the multidisciplinary nature of rhythm, we find different definitions of the term that describe its occurrence in the different aspects of life. However, the study adheres to its meaning as it is used in speech. As such, the term is used throughout the study to refer only to speech rhythm. Abercrombie (1967, p. 96) defined it as ‘periodic recurrence of some sort of movement, producing an expectation that the regularity of succession will continue’. Thus, rhythm is that 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, p. 245).

1.10.3 Stress-Timed Rhythm

The pattern produced in speech can be a result of the regular occurrence of a prominent speech element like stress, which is typically found in a language like English. In other words, the stress tends to be isochronous in the language, appearing at regular intervals. In such cases, the rhythm of the language is described as being characterized with stress-timing features which reflect the heavy stresses and considerable variations in the length of the vowels in the language, (Ladefoged, 2006, p. 246).

1.10.4 Syllable-Timed Rhythm

The regular succession of features of speech can be in terms of syllables, which is a case found in a language like Spanish. As such, it is the syllables of the language, not the stress,

that tend to be isochronous and occur at regular intervals. The languages that display such type of syllable isochrony are called syllable-timed languages which have relatively constant vowel lengths, (Ladefoged, 2006, p. 247).

1.10.5 Rhythmic Indexes

The application of rhythmic indexes or indexes is a new approach to the investigation of rhythm. They are formulae used to measure the durations between intervals of vowels and consonants providing the degree of variation that these measurements display. These indexes are according to Ramus *et al.* (1999) used in order to interpret the impressionistic account of stress-timed and syllable-timed rhythm through particular structural features in the language, namely, syllable structure and vowel reduction, which are said to be affecting the durations of vowels and consonants.

1.11 Outline of the Chapters

Chapter 1 introduced the background of the study and the statement of the problem. The purpose and objectives are presented and the research questions are stated in this chapter. It also proceeds with the scope and limitations and the significance of the study and finally provides definitions of key terms in the study.

Chapter 2 presents a description of Arabic language and its dialects by explaining the features of SA compared to both YAD and JAD. This chapter also accounts for the concepts of prosodic and suprasegmental features in general and rhythm in particular. The chapter presents the different approaches to speech rhythm and proceeds with a review of

the previous studies that have to do with rhythm both globally and in Arabic. The theoretical and analytical frameworks of the study are presented in this chapter. Some of the most prominent rhythmic indexes are also reviewed.

Chapter 3 explains the research design and provides a description of the materials used for data collection and the background of the speakers who participated in the study. The methods of data collection and analysis are also clearly explained.

Chapter 4 reveals the findings of the study that are produced by the discussion and analysis of the data. This chapter presents tables that illustrate the values of measurements that are produced from the speakers' different utterances and compares between them by means of a *t*-test to find out whether or not, there is a significant difference between the rhythmic structures of both Arabic dialects.

Chapter 5 sums up the findings of the study referring to the research questions and objectives to find out whether the questions have been answered and the objectives have been fulfilled.

1.12 Conclusion

This chapter introduced the background of the study and presented a brief description of Yemen and Jordan where the dialects under investigation are used. The statement of the

problem, the purpose and objectives of the study and the research questions are also stated in this chapter. It proceeded with the scope and limitations and significance of the study, and finally presented definitions of the terms involved in the research. The next chapter will present a description of the key features of SA such as the syllable structure, the vowels and the consonants, and compares them to those of YAD and JAD. It will also present a review of prosody in general and rhythm in particular as well as reviewing the previous studies that dealt with rhythm. The theoretical and analytical frameworks of the study are also presented in the following chapter.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents a description of Arabic language by describing the syllable structure, vowels and consonants of SA, YAD and JAD. It also presents an explanation of the concepts of prosody and identifies some of the prosodic features of human speech like intonation, tone, stress and rhythm. Different approaches to rhythm will be discussed starting from the perceptual analyses of rhythm to the approach of isochrony and finally rhythm typology. The chapter progresses with a discussion of the different rhythmic indexes that have been proposed by different researchers since the late 1990s as a current approach of speech rhythm. Some of the studies that have investigated the rhythm of Arabic language are also presented. The theoretical and analytical frameworks of the study are also provided in this chapter.

2.2 Arabic Language and its Dialects

Arabic language is a member of the Semitic language family which is according to Hamdani, Selouani and Boudraa (2010), is spoken by about 300 million people around the world. Apart from being the language of the Holy Qur'an and the means of Muslims' religious practices, it is the official language in all Arabic speaking countries. According to (Watson, 2002, p. 17) Arabic is the common official language in twenty countries that

range from Western Asia to North Africa. These countries include Saudi Arabia, Qatar, Kuwait, Bahrain, the United Arab Emirates, Yemen, Oman, Jordan, Syria, Iraq, Lebanon, Gaza, the occupied west Bank in Palestine, Egypt, Tunisia, Algeria, Morocco, Libya, Sudan, Djibouti, Somalia and Mauritania. Like the other Semitic languages Arabic, is characterized by a limited vocalic system and a rich consonantal system. It is mainly based on the system of consonantal root (Awde and Samano, 1986, p. 15). That is to say, Arabic words are derived usually from a three-letter root which is mainly a verb. This root undergoes processes of affixation changing it to other words. Overall, Arabic has twenty-eight consonants with probably nine places of articulation, (Watson, 2002, p. 46). Arabic has three short vowels, namely, /a/⁴, /i/ and /u/ and three long ones, namely, /aa/, /ij/ and /uw/ and they appear in the form of diacritic marks above or below the consonants, (Al-Shuaibi, 2010). In terms of the writing system, the letters are written from right to left without any capitalization and are connected to each other with the exception of six letters, namely, أ /ʔ/, د /d/, ذ /ð/, ر /r/, ز /z/ and و /w/ which are written separately in Arabic scripts if they occur in the beginning of a word.

The terms Classical Arabic and modern SA refer to the medieval and modern varieties of Arabic. The Classical Arabic is the language of the Holy Qur'an which is quite distinct from the stylistic and lexical elements of modern SA. However, both varieties display no differences in terms of morphology and syntax which remained invariant thus far, (Fischer, 1997, p. 188). Conversely, the vernacular Arabic dialects have remarkably changed over the centuries, and although SA has remained the standard variety in all the Arabic speaking

⁴ All the phonetic symbols used in describing the Arabic language are adopted from the work of Thelwall and Sa'adeddin (2003, pp. 51-54) of describing the Arabic sounds in the Handbook of the IPA.

communities, it ultimately became associated with a great number of regional dialects. Thus, SA cannot be considered as the mother tongue of any Arab child. Rather, a child is brought up speaking his regional variety of Arabic, while SA is learnt as part of the child's education in school. SA today is used in formal written form, while in the spoken form, its use is limited to news reading that is broad cast. The regional varieties of Arabic, on the other hand, are used in all other situations even at formal occasions.

In describing the situation of Arabic, Bedawi (1973), Schmidt (1974) and Harry (1996) suggested that although SA is the official language in all the Arabic speaking countries, it is restricted to very limited contexts or mainly to written texts leading to the prevalence of the colloquial Arabic dialects in the different Arabic regions. Thus, due to the dominant use of the colloquial dialects and the restriction of SA in the written form, these authors suggested that, with time, these dialects are capable of replacing the SA by arguing that the Arabic dialects are used in different formal and informal situations. As such, Blanc (1960), Bedawi (1973), Schmidt (1974), Meiseles (1980) and Harry (1996) treated the situation of the Arabic speaking countries as involving: (i) a standard form (SA which is limited to written texts), (ii) colloquial of the educated, (iii) colloquial of the enlightened (partially educated) and colloquial of the illiterate. They considered the 'colloquial of the educated' as being a semi-literary and elevated form of the language that is used even in very formal situations. In other words, it is the formal form of the dialect that is used by educated people in the Arabic speaking communities. As such, apart from SA, all the Arabic dialects, including YAD and JAD which are the dialects under investigation, will have a colloquial form that is used by the educated; a colloquial form used by the enlightened (partially educated) and a colloquial form used by the illiterate.

The different dialects of Arabic can roughly be divided into two main geographical areas: the first is represented by eastern dialects spoken in western Asia along with Egypt and Sudan; the second is represented by western dialects spoken in countries of the northern part of Africa. The main phonological features which distinguish the western dialects include reduction of short and unstressed vowels which result in consonant clusters even in word initial position, a case which is not common in eastern dialects. For example, the Arabic word *samin* ‘fat’ (Elmedlaoui, 1995, p. 139) is pronounced as ‘*smin*’ with a reduction of the unstressed short vowel /a/ in the western Arabic dialects leading to an initial two consonant cluster which is a feature that cannot occur in SA nor eastern Arabic dialects. Watson (2002, p. 20) pointed out another difference in terms of stress, where the western dialects display iambic word stress-system that contrasts with the trochaic word-stress system of the eastern dialects. Accordingly, an Arabic word like *daras* ‘he studied’ would be stressed as /da'ras/ in western dialects, but /'daras/ in eastern dialects.

YAD and JAD, which are the dialects under investigation in this study, are Arabic dialects spoken in Yemen and Jordan, respectively. Whereas JAD is one of the dialects of the Levant, YAD is one of the dialects of the Arabian Peninsula, which according to Watson (2002, p. 21) has received little attention in phonological and morphological studies of Arabic. The YAD is close to SA in terms of phonology since it has maintained much of the features of its inventory. It has undergone less linguistic changes than other Arabic dialects which display great variations in terms of consonant pronunciations, (Watson, 2002, p. 27). Some of the most notable changes that these dialects display are the substitutions of the sounds /θ/, /ð/, /ð^s/, and /q/ with /t/, /z/, /z̤/ and /ʔ/, respectively, in Egyptian, Lebanese, and Jordanian and Syrian dialects. Similar changes can be recognized in Sudanese dialect with

the only exception of the sound /q/ which, unlike the previous dialects, is substituted with /ɣ/ instead of /ʔ/. For instance, the SA word *qal* ‘said’ is pronounced as /ʔal/ in Egyptian Arabic dialect while in Sudanese Arabic dialect, it is pronounced as /ɣal/.

2.2.1 Syllable Structure of SA, YAD and JAD

English syllables can range from one to four consonants. (Laver, 1994, p. 33) mentioned that English can have sequences of three consonant clusters in the initial and middle position of a word and can have four consonant clusters in word final position. However, in SA, the syllable structure can have the following types:

- cv *fi* ‘in’
- cvc *dar* ‘house’
- cvcc *kenz* ‘treasure’

Unlike English structures, the above structures show that in SA, syllable onsets can never be vowels, syllables have always to begin with a consonant. Moreover, two consonants cannot meet in the beginning of the syllable but they are allowed to meet in syllable codas.

As for the YAD, it has been treated by researchers like Al-Shuaibi (2009), Jalal (2011) and Na’ama (2011) as having a syllable structure similar to that of SA. However Watson (2002), argued that YAD has an additional **cvccc** syllable structure which occurs depending on the type of word-final consonants. That is to say, the **cvccc** syllable structure can occur in YAD only if the word involves a final /t+f/ which is a suffix that denotes a 1st person

singular subject or a 2nd second person masculine singular plus the negation marker. For example:

a. ma-akala-t-sh /maʔakalatʃ/

no-ate-she-not

She didn't eat.

b. ma-katab-t-sh /makatabtʃ/

no-wrote-you-mascular-singular-not

You didn't write.

Watson (2002, p. 57) suggested that the final /tʃ/ sound is perceived as an affricate or a single consonant in YAD despite the fact that it is derived from two different morphemes, namely, the pronoun morpheme and the morpheme of negation.

Like YAD, JAD also retained the three syllable structure of SA. However, Al-Saidat (2010) argued that JAD has a fourth syllable structure that allows for two consonant clusters in syllable onsets which is not common in both SA and YAD. This **ccvc** structure can be seen in JAD in words like '*fribna*' (we drank) and '*smiʃit*' (I heard) which are both pronounced in SA and YAD with the insertion of the vowel 'a' in between the two initial consonants of the words.

2.2.2 Vowels of SA, YAD and JAD

The Arabic language has an impoverished vocalic system compared to its consonants. There are three short vowels in SA, namely, /i/, /u/ and /a/. These vowels, according to Dickins, (1998), can combine with identical vowels resulting in a longer form of the vowel that takes more or less double the time needed for producing the short vowel (Al-Ani, 1970). These long vowels are, as put by Thelwall and Sa'addedin (2003), are /ij/. /uw/ and /aa/. The language also allows for combinations of different vowels resulting in the diphthongs /aw/ and /ay/, (Dickins, 1998, p. 6).

One of the first acoustic studies of SA vowels was that of al-Ani (1970). However, this study as well as other following studies did not seem realistic in the sense that they treated the long vowels in Arabic as having precisely double the duration of the short vowels. In studying the vowels of Arabic, Newman and Verhoeven (2002), argued that SA vowels, particularly /a/ and /aa/, display a high degree of stability, regardless of the dialectical background of the speaker. Thelwall and Sa'addedin (2003, p. 52) presented a description of the Arabic vowels within the framework of the Cardinal Vowels:

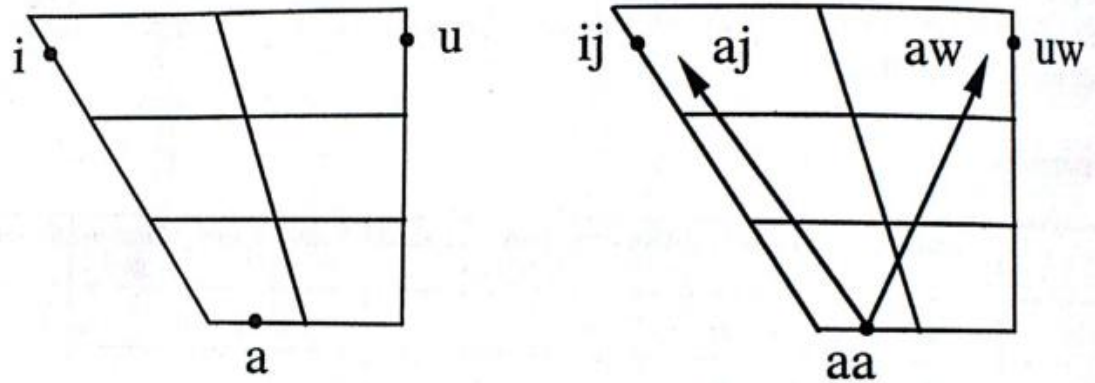


Figure 2.1: The Cardinal Vowel Chart of SA (Adapted from Thelwall and Sa'addedin (2003, p. 52)

Like SA, YAD has three short vowels, namely, /i/, /u/ and /a/. Whereas the /i/ and /u/ were fused over time to a schwa-like vowel in other Arabic dialects which turned out to have a two-short vowel system, YAD still maintains the distinction between these two close vowels. Unlike the short vowels, the three long vowels /ij/, /uw/ and /aa/ are maintained in all Arabic dialects in general and in YAD in particular. However, there are two additional long vowels that appeared in YAD over time, which are slightly open forms of the vowels /ij/ and /uw/. They are represented by (El-Isa, 1982) as /e:/ and /o:/. The two SA diphthongs /aw/ and /ay/ are also preserved in YAD and other dialects of the Peninsula.

Like many other dialects and languages of the world, vowels in YAD are also subject to certain speech processes such as lengthening, reduction or even deletion. In their study of vowel deletion in YAD, Al-Yaari, Al-Hammadi and Luwa, (2012) found that YAD, like several other Arabic dialects, display a phonological feature of vowel deletion. This deletion is caused by syntactic, morphological and phonological rules. Short vowels are often deleted in YAD in past tense verbs. For example the SA word *shariba* 'he drank' is

pronounced in YAD as *shirib* with changing the first vowel from /a/ to /i/ and deleting the final /a/ vowel. Moreover, the SA word *fahimna* (3rd.plu., female) ‘they understood’, is pronounced in YAD as *fi~~h~~mayn* where the first vowel is changed from /a/ to /i/ and the second vowel is deleted while the consonant cluster (mn) is broken by a diphthong /ay/.

As for JAD, the short vowels and their long counterparts are also preserved. However, the SA diphthongs disappeared from the dialect and turned out to be replaced by the long vowels /e:/ and /o:/ as can be seen in the following examples:

SA	JAD
/ʕayn/ ‘eye’	/ʕe:n/ ‘eye’
/mawt/ ‘death’	/mo:t/ ‘death’

Thus, unlike YAD which adopted these two sounds as additional vowels, JAD used them as a replacement of the SA diphthongs.

2.2.3 Consonants of SA, YAD and JAD

The Arabic language is quite rich in its consonantal system and it probably needs a complete study to describe its features. Although it might not be possible to discuss the Arabic consonant sounds in detail in this study, it is quite useful to present the terms ‘emphatic’ and ‘non-emphatic’ consonants. The emphatic sounds are /t^ʕ/, /ð^ʕ/, /s^ʕ/, /d^ʕ/, /l^ʕ/ and /ʕ/, while the remaining sounds are considered as non-emphatic consonants. Abdo (1969) as cited by (El-Isa, 1982) provided a complicated description of emphasis, treating it as a prosodic feature or a function of the syllable rather than isolated segments. As such, the

information given in a particular segment of speech is not a result of single segment, rather, it is a result of the accompanying sounds in that context; which means that in the occurrence of ‘emphasis’ in a particular syllable, the smallest unit of speech would be that entire syllable. However, Ladefoged, (1973) and Ghazali (1977) presented a simplified description of emphatic consonants treating them as pharyngealized sounds which, according to Ghazali (1977) differ from the pharyngeal sounds in that the constriction in the former is in the upper pharynx while the constriction in the latter is below the epiglottis. Following Ladefoged and Ghazali, the current study, treats the emphatic sounds as phonetic segments which display characteristics of pharyngealization rather than prosodic features or functions of the syllable.

Table 2.1 shows the entire inventory of the SA consonants based on their place and manner of articulation, adopted from Thelwall and Sa’addedin (2003, p. 51):

Table 2.1: The SA Consonants (Adapted from Thelwall and Sa’addedin (2003, p. 51)

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	b		t d				k	q		ʔ
Nasal	m		n							
Fricative		f	θ ð	s z	ʃ		x ɣ		ħ	h
Affricate					dʒ					
Trill				r						
Approximant						j	w			
Lateral Approximant					l					

Pharyngealized consonants: **t^ʕ, d^ʕ, s^ʕ, ð^ʕ, l^ʕ, ʔ^ʕ**

The YAD retained all the consonantal system of SA with the only exception of the affricate sound /dʒ/ which in some areas of Yemen is replaced with the stop sound /g/. As for JAD, we find more changes in the pronunciation of the consonant sounds. For instance, the SA consonants /θ/, /ð/, /ðˤ/, and /q/ are replaced with /t/, /z/, /z̤/ and /g/, respectively.

2.3 Prosody

Speech signal according to Miller (1978, p. 175), carries two types of information at the same time. One of them is the isolated phonetic sounds, namely, vowels and consonants, and the other is the prosodic features of speech represented by such features as pitch, loudness and intonation. According to Fujisaki (1997, p. 28) prosody is characterized by both measurable aspects and underlying principles which can only be recognized in a series of cohesive sounds as those found in human speech. As such, prosody can be defined as ‘the systematic organization of various linguistic units into an utterance or a coherent group of utterances in the process of speech production’, (1997, p. 28). Kent and Read (2002) defined the term as ‘features and modifications whose effects transcend the boundaries of individual phonetic elements... and are superimposed on phonetic sequences, giving these sequences a coherence and unity that obscures the ostensible discreteness of these phonetic constituents’, (2002, p. 223). As such, speech is regarded as a string of phonetic elements, namely, segments, produced within the larger scope of intonation, stress, rhythm, loudness and rate, namely, suprasegmentals. Thus, these features are larger than segments and refer to stretches of speech beyond the segment boundaries, (Kent and Read, 2002, p. 227).

Clark *et al.*(2007, p. 326) suggested that the prosodic or suprasegmental features are distinguished from the segmental features, which refer to the discrete sounds of vowels and consonants, in that they are entities and features that cover and extend over larger stretches of speech such as pitch, rhythm and tempo. This distinction can be manifested by the different writing systems that are used to indicate the segmental sounds such as the alphabet of English, French and Arabic, or the non-alphabetic script of Chinese. Suprasegmentals, on the other hand, do not have similar indications. However, the authors argue against the idea of suprasegmentals as being superimposed on a message conveyed by the sounds of vowels and consonants, suggesting that suprasegmentals are completely meaningful entities that contribute to the message being conveyed. As such, they are indispensable elements in the message and an ‘integral part of speech production’, (Clark, 2007, p. 327). Accordingly, prosody provides essential functions in speech and is regarded as an important factor of intelligibility in communication. Speakers usually do not produce uniformly measured monotonous or robot-like utterances. The prosodic or suprasegmental information conveyed by the speaker helps the listener in perceiving what is produced as a complicated stream of continuous speech.

Although this study follows authors like Ball and Rahilly (1999), Crystal (2003) and Clark *et al.* (2007) in treating the terms prosody and suprasegmentals as being synonymous, some researchers use them to refer to two different, though related, concepts. As mentioned earlier in section 1.10.1, researchers like Fox (2000) and Trask (2007), treated suprasegmentals as a phonological concept that is not recognized in terms of segmental features such as the quality of the voice or the state of the vocal tract, while prosody is a

phonetic concept that refers to such features as intonation, accent and length that can be recognized in terms of both segmental and suprasegmental features.

In describing prosody, Clark *et al.* (2007, p. 327) pointed out that one should take into account a series of effects that range along a continuum from linguistic to extralinguistic effects of the information carried by the speech signal. That is to say, at one extreme of the continuum lies the linguistic effects represented by features like stress, tone and intonation which vary from language to language in their organization; and at the other extreme lies the extralinguistic or non-linguistic effects such as the voice quality which results from the nature of the larynx and vocal tract that vary from speaker to speaker. In between these two extremes lies the paralinguistic effects which can be represented by such factors as nervousness or fatigue.

There are a number of suprasegmental features in human utterances such as stress, tone, rhythm and intonation. Intonation is as Ladd put it ‘the use of suprasegmental phonetic features to convey ‘postlexical’ or sentence-level pragmatic meanings in a linguistically structured way’, (Ladd, 1996, p. 6). Pike (1945, p. 21) described intonation as a feature that is superimposed on the essential meaning of words in order to convey the attitude of the speaker rather than carrying a fundamental meaning. Ladefoged defined intonation as ‘the use of pitch variations to convey syntactic information’, (Ladefoged, 2006, p. 248). However, whereas the pitch variations occur in the level of the phrase in intonation, in tone, the pitch variations occur in the level of a word to convey both the meaning and the grammatical function of the word. This is especially true in a language like Chinese where

special patterns of pitch are included within syllables or words to convey different meanings even if the remaining phonetic content of the syllable or the word is not changed, (Clark *et al.* 2007, p. 338).

Stress is also one of the suprasegmental features of human speech which causes the syllables in which it falls to be more salient in the flow of speech and produced with greater energy than those which are unstressed. Adams (1979, p. 58) cited Jones' (1932) description of stress as a 'degree of force with which a sound or syllable is uttered', and that, it is a subjective action that includes a strong push of the wall of the chest producing an 'objective impression of loudness'. This definition does not only include the articulatory or muscular energetic actions, but also underscores the role of the speaker and listener who produce and perceive this process.

Stress involves the expansion of muscular energy and the expulsion of air from the lungs as a result of the rib cage contraction and perhaps an increase in the pitch that might lead to a greater length of the sound produced, (Ladefoged, 2006, p.243). Many researchers describe stress merely in terms of loudness, however, Ladefoged suggested that this might be a misleading oversimplification given that loudness is associated with acoustic energy and some sounds might be characterized by high acoustic energy due to different degrees of mouth opening regardless of stress, (Ladefoged, 2006, p. 243).

Differences in the use of stress in different languages result in different rhythms in these languages. Although there is a strong relation between stress and rhythm given that the former serves as an indication of the rhythmic impulse, stress is only one factor that can influence rhythm.

Rhythm is one of the components of the prosody of human language. It is described by Handel (1989, p. 384) as an experience that ‘involves movement, regularity, grouping, and yet accentuation and differentiation’. Gut (2012, p. 83) referred to the term as a ‘temporal organization of a language. However, the term has been described in different ways and through different approaches which are going to be discussed in the following section.

2.4 Speech Rhythm

Rhythm is a multidisciplinary topic that is manifested in different fields that range from engineering, biology, medicine, psychology, neuroscience and computer science to philosophy, literature, music, and linguistics. Accordingly, we find different definitions of the term that describe its occurrence in the different aspects of life. A general definition is found in *Oxford English Dictionary* (1989) describing rhythm as ‘a regulated succession of strong and weak elements’ which conforms to Sonnenschein’s (1925, p. 16) definition of rhythm that refers to the term as a succession of events in time. Regardless of the different meanings of rhythm, the term in this study refers to speech rhythm which is generally perceived as ‘movement’, (Fry, 1964, p. 217). According to Ball and Ranhilly, ‘it is usually thought that there is some sort of rhythmic structure that underpins speech production and perception’, Ball and Ranhilly (1999, p. 120). Crystal (1985, p. 266-7) defined rhythm as

the regular occurrence of prominent units which are perceived in speech in terms of stressed/unstressed syllables, short/long syllables, or high/low pitch, or even a combination of these features. As such, Crystal put forward that the properties which contribute to rhythm are pitch, loudness and tempo, (Crystal, 2007, p. 171). Thus it is recognized by a number of features that vary with time and situation variation, (Keller and Zellner, 2000). Abercrombie (1967) adheres to the concept of ‘movement’ and ‘regularity’ in his description of rhythm defining it as a movement that occurs repeatedly creating an anticipation of the continuity of this regular succession. Gibbon and Gut (2001, p. 91) defined rhythm as ‘the recurrence of a perceivable temporal patterning of strongly marked (focal) values and weakly marked (non-focal) values of some parameter as constituents of a tendentially constant temporal domain (environment)’. The terms ‘focal’ and ‘non-focal’ values are used to refer to sequences like high/low pitch, long/short syllables and segments of vowel and consonants.

As a matter of fact, all human speech displays rhythm (Abercrombie, 1967), which is recognized as a rhythm of movement that occurs in a stream of continuous speech as some kind of a regular sequence of movement or periodic recurrence of a kind of movement (Abercrombie, 1967). This movement mainly involves stress and syllables, and is experienced by both the speaker and the listener. The recurrence of these two elements of speech results in what has been described by Pike (1945) as stress-timed or syllable-timed rhythm. According to (Abercrombie, 1967, p. 36), the syllable process and the stress process are the basis on which the entire speech is built and they are used in different ways in different languages depending on the way in which stressed and unstressed syllables follow each other to produce what is known as the rhythm of a language. In explaining

these processes, Abercrombie (1967) refers to speech rhythm as a muscular rhythm and these muscles are the breathing muscles. It is the recurrence of the ‘chest-pulses and stress-pulses’ that determines the rhythm of a language. As such rhythm is originally a ‘muscular rhythm; a rhythm of bodily movements rather than a rhythm of sound’, (Abercrombie, 1967, p. 19).

Based on these definitions and explanation, one can infer that rhythm is that prosodic feature of human speech that results from isochrony, namely, the regular occurrence of units of speech like syllables or stress which according to Laver (1994) characterizes all human languages. Thus, rhythm is sensitive to the successive nature of features, like syllables and stress, which helps to classify languages into, what Pike (1945) referred to as stress-timed and syllable-timed languages.

Stress-timed rhythm displays considerable variations in terms of syllabic durations and structures produced by ‘the stress-pulses’ process as can be recognized in languages like English, Russian, Arabic and German. In other words, in these languages the stressed syllables recur at equal intervals (they are synchronous) (Abercrombie, 1967). The syllable-timed rhythm, on the other hand, shows less variation in terms of syllabic duration and more regularity in syllabic structures, (Keller and Zellner 2000). That is to say, the syllables occur at equal intervals (they tend to be isochronous) produced by ‘chest-pulses’ process as in languages like French, Spanish and Telugu (Abercrombie, 1967, p. 97). However, when Bolinger (1965), Duckworth (1967), O’Connor (1965), Shen and Peterson (1962) measured

the durations of inter-stress in English, they could hardly find any evidence of physical isochrony of stresses as Abercrombie (1967) argued, (Cumming, 2010).

Bloch (1950) found that Japanese rhythm can be regarded as isochronous not in terms of stress or syllables, rather in terms of morae which recur at regular intervals in Japanese. Grabe and Low (2002, p. 515) defined morae as subunits of syllables that consist of a short vowel and an onset consonant that precedes that vowel. Later researches, therefore, treated mora-timing rhythm as a different type of speech rhythm which is best exemplified in Japanese.

2.4.1 Different Approaches to Rhythm

Generally, speech rhythm is not a new concept. Aristotle describes the term rhythm as a recurrent pattern found in speech or other sounds, (Cumming, 2010). Even in the 1900s some psychologists were interested in studying how listeners perceive rhythm. Wallin (1901) for instance measured the feet duration in reading English poetry and found that it displays rhythmical lines despite the fact that it possesses unequal durations of feet. Brown (1911) on the other hand finds that whereas verse rhythm is irregular, it displays considerable regularity in terms of recurrence. Later studies started to move from the subjective explanations and descriptions of rhythm and merely impressionistic findings, towards a more phonetic perspective by conducting several experiments, recording speech sounds and using some modest tools for measurement. English, German and French were the most languages that have been investigated by researchers.

However, the nature of ‘movement’ that characterizes rhythm has been perceived differently by researchers over time. One of the earliest approaches to rhythm that prevailed until the early eighteenth century is the quantitative approach that viewed rhythm as a result of series of long and short syllables that occur in speech, (Adams, 1979, p. 56). Based on this approach, English was described in terms of Greek and Latin, the fact which was first supported by Herries (1773) who found that introducing the intervals, accents and pauses that appear in Greek and Latin into English is not at all problematic or inconvenient. However, studies like Omond (1921), Thomson (1926), Guest (1938) and Sumera (1968) opposed these practices arguing that the feet and rules which control the quantity both in Greek and Latin are inapplicable in the case of English. As such, applying this approach to a language like English was not favoured for the lack of similarities between English and those languages, which appear to be incompatible with the assumption that claims a possible comparability between them, (El-Isa, 1982, p. 81). Moreover, this approach seemed to fail in considering the phonetic aspects of the language. In arguing against the quantitative approach, Thomson (1926, p. 28) presented his argument by suggesting that one cannot count merely on syllable quantity in perceiving rhythm and that quantity and accent are inseparable in much the same way as the environment and inheritance are.

Other theories appeared in the analysis of rhythm such as the non-temporal approach that became a traditional view in the nineteenth century. This approach ‘postulates accentual rhythm’, (Adams, 1979, p. 56), in other words, rhythm is determined by the regular occurrence of accent or stress. As such, it is the stress or accent, not the syllable, that defines the rhythmic patterns in a language. In referring to verse, Barkas (1934) further subdivides this theory into (i) the isosyllabic theory, which assumes that the equality in a

line is a result of the equality in the number of syllables, (ii) the accentual theory which assumes that the equality in a line is a result of the equality in the number of accentual syllables, and (iii) the syllabic-accentual foot theory which assumes that the equality in a line is a result of the equality in the number of groups of both accented and unaccented syllables.

Guest (1838), Draat (1909) and Chatman (1965) argued in favour of the non-temporal approach by saying that accent is the only factor that determine rhythm. Sweet (1913, pp. 11-12) also stated that there is no question that the inclination to alternate between weak and strong stress is the origin of rhythm. As such, it is the regular exploitation of these natural tendencies of alternating between weak and strong stress that govern rhythm.

Due to the failure of these approaches in presenting a comprehensive description and efficient analysis of rhythm, a new approach, known as the temporal approach, emerged in an attempt to succeed in what the previous approaches failed to achieve. According to this theory, neither syllable quantity nor accent alone can govern rhythm. It is time that is regarded as the most important factor that determines rhythm of a language. Steel (1775, p. 68), suggested that rhythm is determined by periodic pulsation. This view was later on embraced by Chapman (1818) and Patmore (1857, p. 224). However, Steel (1775) argued that despite the periodic pulsation in a language like English, monotony is not recognized due to the recurrence of pauses and considerable variations in the syllable durations. His observations coincide with subsequent arguments that regarded silence and pauses as important features of English rhythm referring to such features as ‘silent stress’, (Adams, 1979, p. 27). However, the claim of periodic pulsation postulated in this approach appeared

to many researchers as untenable due to the lack of objective measurements that support the theory when it was presented, and the failure to provide ways of reliable analysis that may result in some empirical evidence of the theory, (Adams, 1979, p. 56).

In the early decades of the twentieth century, the temporal approach began to be presented within the concept of isochrony in speech which postulates the regular occurrence of elements in speech. Both Pike (1945) and Abercrombie (1967) agree that rhythm is primarily dependent upon the isochronous intervals in speech. Jones (1960) and O'Connor (1968) described the concept as the occurrence of stressed syllables in almost equal intervals. Classe (1939, p. 85) suggested that isochrony occurs in particular conditions where the number of syllables in the groups concerned and the phonetic and grammatical structures of these syllables do not vary widely. In his findings of what is referred to as 'syllable-time' and 'stress-timed' speech, Classe argued that the former displayed approximately equal recurrent intervals in terms of syllables, while in the latter, it is the stress that occurred at approximate equal intervals. Although these two types of rhythm have been referred to by Lloyd James (1940) as 'machine-gun' and 'Morse-code', the use of these terms has faded when Pike (1945) presented the terms 'syllable-timed' and 'stress-timed', respectively.

With regard to English rhythm, Pike (1945) explained that it is the stresses that occur at equal intervals that produce the stress-timed rhythm of the language. In a language like Spanish, on the other hand, it is the syllables that occur at equal intervals with very little chances of vowel reduction or modification producing the syllable-timed rhythm of the language. Therefore, when a Spanish speaker speaks English, he/she has to abandon his

‘sharp-cut syllable-by-syllable pronunciation’ (Pike, 1945, p. 35) and try to reduce or lengthen vowels where necessary to adapt to the stress-timed rhythm of English. Most of the researchers at that time dealt with languages like English, German and French producing comparable results in that the first two display isochrony in terms of stress while French possesses durationally equal syllables.

Generally, many researchers, like Wenk and Wioland (1982), argued that rhythm cannot be regarded as primarily dependent upon the isochronous intervals in speech. They suggested that this oversimplification might be unreliable due to the fact that their studies found that the French syllables are not isochronous. Dauer (1983) also found that English syllables are no more isochronous than the other seeming syllable-timed languages Italian, Spanish and Greek. Probably, only Japanese can be considered as been isochronous since its morae tend to occur at equal intervals (Port, Dalby, O'Dell, 1987). In studying the American English speech, Nakatani, O'Connor and Aston (1981) could not find evidence that support isochrony. The duration of syllables in their study seemed to increase with the size of the word, which is not compatible with the concept of isochrony that postulates a faster production of longer words than the shorter ones, (1981, p. 103). Deterding also added that perfect isochrony does not exist in languages and that a complete regular repetition of speech units would sound more ‘like a robot’ than natural human speech, (1994, p. 316). Hoequist suggested that this concept of syllable/stress-timed isochrony might be unreliable unless a non-strict interpretation of the term is presented, (1983b, p. 229).

Despite the different reactions of researchers towards the concept of isochrony, the remarks of Pike's (1945) about the stress-timed and syllable-timed rhythm, followed by those of

Abercrombie (1967) laid the foundation of a different approach proposing that languages can be discriminated based on the stress-timing or syllable-timing features of their rhythm. Researchers, therefore, started investigating the rhythm of different languages to examine the accuracy of this assumption.

Despite proposing that languages can be classified based on the features of their rhythmic structures, Pike (1945) argued that a language can display different kinds of rhythm depending on the style of speech. In English for instance, although it is rare, yet, there can be cases where rhythm depends on the prominence of the syllables recurrence rather than stress. However, Abercrombie (1967) disagreed with the idea of a multiple kind of rhythm in one language, for according to him, although there is no obvious reason, there is only one type of isochronous succession in a language which if occurs, the other will not, (Abercrombie, 1967, p. 97). Dasher and Bolinger (1982) however, suggested that stress-timed and syllable-timed properties can be correlated in languages depending on their phonological structures. In his perceptual study of rhythm, Miller (1984) found that both the phoneticians and untrained listeners, who participated in the study, were inconsistent in identifying stress-timed or syllable-timed languages from hearing both read and naturally spoken sentences. Based on those results, he argued that two types of rhythm do exist in one language, which makes it difficult for the listener to make a decisive classification of a language. However, he did acknowledge that an important relevant factor leading to these results may be the prejudice of the phoneticians to their training backgrounds and the lack of experience of the untrained listeners.

In arguing against this rhythm typology, Balasubramanian (1980) argued that Tamil has a kind of rhythm that is neither syllable-timed nor stress-timed. Syllables in certain structures in Tamil display regularity and thus expected durations. Pointon (1980) also remarked that although, based on Pike's stress-timed and syllable-timed classification, Spanish and French are regarded as belonging to the same category, it appeared in his study that Spanish resembled English more than French in that they both display contrastive stresses. Thus, in dealing with the rhythmic structures of these languages, it is more plausible to consider English and Spanish as belonging to the same category rather than having Spanish and French classified as syllable-timed languages and English as a stress-timed language, (Pointon, 1980, p. 293).

2.4.2 Dauer's (1983) Approach as a Theoretical Framework of the Study

Later studies have argued against Pike's and Abercrombie's assumption that emphasized the dichotomous nature of rhythm which placed languages at binary oppositions as either stressed-timed or syllable-timed languages. These studies have revealed the problematic nature of this dichotomy with the complications that emerged from classifying languages like Catalan and Polish due to the mixed features of their rhythmic structures. As such, a new approach has emerged in the 1980s that studied rhythm in terms of a reflection of certain phonological features that result in rhythmic diversity which cannot be dichotomous. This approach, upheld by researchers like Dasher and Bolinger (1982), Dauer (1983, 1987) and Nespor (1990), considered rhythm to be determined by the absence or presence of certain phonological and phonetic features of a language such as the variety and complexity of syllable structure, vowel reduction and lexical stress, Dauer (1983). In

studying the inter-stress intervals in English, Thai, Spanish, Italian and Greek, Dauer (1983) found that the stress recurs more or less regularly in a prototypical stressed –timed language like English and a prototypical syllable-timed language like Spanish. As such, the regular recurrence of stress seems to be a universal property in languages and the difference between these two languages that have been considered as belonging to two different classes is not caused by the durational intervals of interstresses in her study. Despite the more or less regular recurrence of stress in both English and Spanish, their rhythm seems to be different. Influenced by the remarks of Classe (1939, p. 132), Dauer suggested that the rhythm output is an automatic consequence of certain linguistic circumstances, (1983, p. 60) or in other words the interaction of a number of factors in a language, namely, the syllable structure, vowel reduction and stress. Therefore, Dauer (1987) proposed that it would be rather practical and objective to perceive stress-timed languages as having complex structure and several cases of vowel reduction, while syllable-timed languages display more constraints and mora-timed languages showing even much more constraints. Tan, (2011, p. 62) also suggested that 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. This would help to deal with languages like Catalan and Polish which seem to be problematic when trying to classify them exclusively to a particular type of rhythm. Catalan, for instance, possesses a syllable structure similar to that of Spanish, yet, it displays cases of vowel reduction. Polish on the other hand possesses a complex syllable structure typical of stressed-timed languages, yet, it displays no vowel reduction. That is why linguists have never been in agreement over the classification of these two languages (Ramus *et al.*, 1999). Dauer (1987), therefore, suggested that languages should not be thought of as dichotomous, in other words,

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, but also others would stand mid-way between the two extremes for displaying mixed features of rhythm.

Nespor (1990) further supported Dauer's argument that the stress and syllable-timed rhythm are the result of a series of certain phonological processes, (1990, p. 158). She used the process of acquisition of phonology as an evidence for this theory. In the early stages of acquisition when the phonological system is not yet completely developed, children of a prototypical stress –timed language like English and children of a prototypical syllable –timed language like Italian would produce utterances with temporal structures that are more similar at this stage than at later stages. In other words, the temporal patterns of two languages that are classified differently in terms of rhythm would sound similar in the early stages of acquisition due to the very uniform structures produced at this stage with absence of the phonological processes that contribute to different patterns of rhythm such as vowel reduction or stress, (1990, p. 165).

In their acoustic study, Ramus *et al.* (1999) suggested that the measurements reflect certain phonological properties that may have an influence over the rhythmic structure of the language under investigation. However, Arvaniti (2012) claimed that although Dauer's criteria were successful in the classification or distribution of prototypical languages, Dauer did not provide testing for the level of consistency of the distribution of other non-prototypical languages along a continuum based on her criteria. Barry and Andreeva (2001)

found that vowel reduction, one of the criteria proposed by Dauer, is not an exclusive feature of stress-timed languages as it can also occur in syllable-timed languages. However, this conclusion was drawn without a distinction between vowel reductions that occurs as a result of fast speech rate and between the reduction of the vowel which is influenced by the local context or the prosodic position of the vowel. If we count on instances of vowel reduction that results from faster rates of speech then all languages would probably be classified as stress-timed languages.

In considering Dauer's criteria, the following phonological features are described

2.4.2.1 Syllable Structure

The syllable structure in stressed-timed languages, as Dauer (1983) posited, display a wide variety of syllable structures than syllable-timed languages which have relatively restricted structures of their syllables. The variety of syllable types results in differences in syllable length. The stress-timed languages also allow for consonant clusters in their syllables while syllable-timed languages do not allow for clustering of consonants and tend to simplify their syllable structures by means of phonological processes such as epenthesis. In the prototypical stress-timed English, we find that the syllables have a nucleus that 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, the prototypical syllable-timed Spanish displays a syllable structure with a nucleus that consists of a vowel or diphthong and the syllable can allow for two consonants only word medially, (Dauer, 1983, p. 55).

According to Dauer (1983) the variety of syllable structures in a language enhances the stress timing features of that language. A stress-timed language like English appeared in her

study to have different types of syllables whereas the syllable-timed Spanish did not display such a variety with the majority of the syllable types were the CV syllable structure:

ENGLISH	SPANISH
CV (34%)	CV (58%)
CVC (30%)	CVC (22%)
VC (15%)	CCV (6%)
V (8%)	V (6%)
CVCC (6%)	

Dauer (1983, p. 56)


2.4.2.2 Vowel Reduction

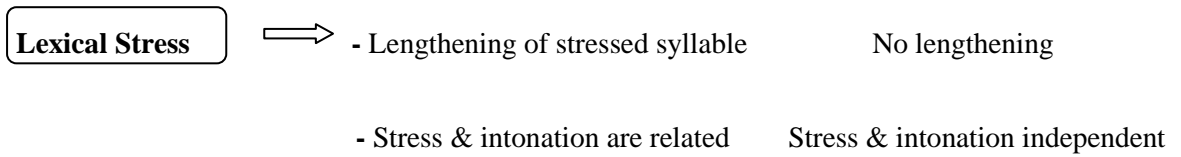
Dauer (1983) described stressed-timed languages as those which tend to reduce the vowels in unstressed syllables resulting in the irregularity of the syllable structure of these languages, whereas syllable-timed languages do not employ vowel reduction resulting in more regular syllable structures. This process of vowel centralization in unstressed syllables is found in languages like English, Swedish and Russian. However, the author mentioned that the vowel reduction in English does not affect the number of syllables, for instance the reduction in ‘saw him’ /sɔ:əm/ does not reduce the number of syllables to one syllable as /sɔ:m/. On the contrary, in French, the reduction in ‘*e muet*’ leads to a reduction in the number of syllables, (1983, p. 57). Syllable-timed languages usually do not involve vowel reduction, however, in Spanish, fast speech might lead to a process of reduction but in terms of consonants rather than vowels; a process which does not affect the length of a syllable, (Dauer, 1983, p. 58).

2.4.2.3 Stress

Stressed-timed languages 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. On the contrary, in syllable-timed languages, there is no evidence of lengthening and intonation is not dependent on stress. Delattre wrote that ‘syllabic stress affects syllabic duration for every position or type of syllable, but this conditioning is relatively strong in English, weak in Spanish and medium in German’ , (1966, p.196). In further supporting this notion, Dauer posited that a language would be more stress-timed if the sequences of interstress intervals do not display considerable variations in the number of syllables. In English, the number of syllables in successive interstress intervals ranges from one to four syllables only whereas in Spanish, the number of syllables can reach up to seven syllables. Table 2.2 summarizes these phonological features following Dauer’s criteria.

Table 2.2: Summary of the phonological features following Dauer’s criteria

	STRESS-TIMED LANGUAGE	SYLLABLE_TIMED LANGUAGE
		
<div style="border: 1px solid black; padding: 2px; display: inline-block;">Syllable Structure</div>	⇒ - Variety of syllable structures - Allows consonant clusters	Limited syllable types Limited consonant clusters
<div style="border: 1px solid black; padding: 2px; display: inline-block;">Vowel</div>	⇒ - Involves vowel reduction	No vowel reduction



Thus, according to this approach, a simple dichotomic classification is misleading. Dauer (1983) viewed the rhythm in terms of a continuum model where languages can be regarded as being more or less stress-timed depending on the phonological features that enhance the stress timing effects in a language, or the extent to which stress plays role in that language. This allows for the description of some languages as occupying an intermediate position in a continuum that ranges from less to more stress-timed rhythmic structure rather than from syllable-timed to stressed-timed rhythm. In her view, the rhythm of all languages is based on stress and the difference between the rhythms of languages like English and French or Spanish lies in that the stressed syllables in the former are more prominent than the latter.

In investigating the rhythmic structures of eighteen languages, Grabe and Low (2002) found that languages range from the stress-timed extreme to the syllable-timed end while others stood mid way between the two extremes for their intermediate features. This supports Dauer's postulation of the continual nature of rhythm where no evidence is found for dichotomous or binary clusters of languages.

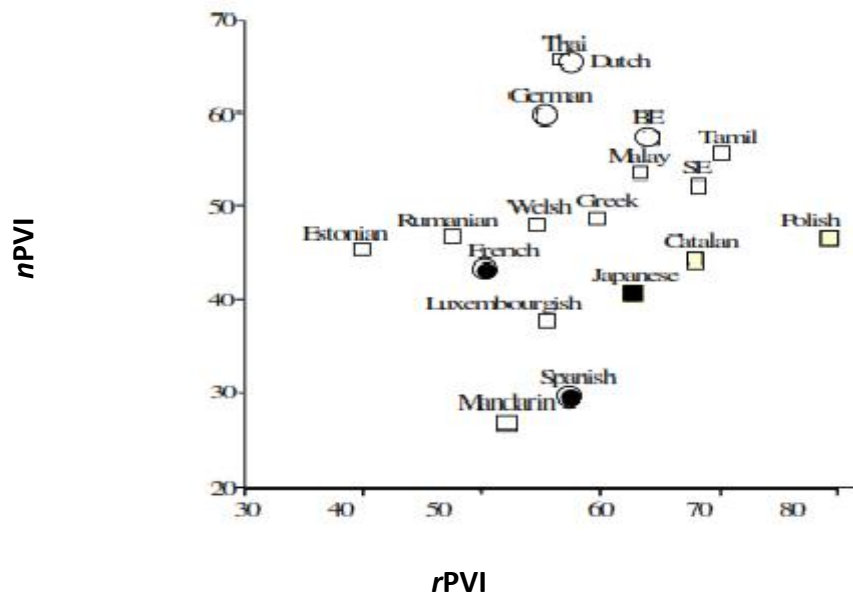


Figure 2.2: Distribution of other languages over the *nPVI* and *rPVI* scales from Grabe and Low (2002, p. 7/16)

However, in investigating the rhythm of Estonian, English, Mexican Spanish, and Castilian Spanish, Nolan and Asu (2009) presented the term ‘coexisting’ rhythm, rather than ‘intermediate rhythm’, as a possible notion that needs to be considered in investigating and describing speech rhythm, (2009, p. 76). This term has been particularly appropriate for describing the case of Estonian which appeared to be rhythmically similar and rhythmically different from English based on a bi-dimensional study of rhythm. Although, in terms of syllables, Estonian displayed more regularity than English, in terms of foot, both languages tended to be similar.

2.5 Rhythmic Indexes

Most of the previous investigations of rhythm were based on auditory perceptions. It was only in the late 1990s that new and advanced instrumentations for measuring rhythm have

been proposed. In recording two to three year old English children, Allen and Hawkins (1979, 1980) found, based on auditory observations, that the older ones tend to exhibit more reduced syllables and stress patterns similar to those of adults. They suggested that a child's rhythm is initially syllable-timed but later can develop into a stressed-time rhythm. However, they acknowledged that this simplification needs more research and investigation. In studying French infants' perception of rhythm, Nazzi, Bertoncini and Mehler (1998) revealed that infants could distinguish between English and Japanese rhythms, as well as distinguishing Italian and Spanish rhythms from those of English and Dutch, however, they could not distinguish between English and Dutch rhythms. Nazzi *et al.* (1998) therefore suggested that rhythms of languages are dependent on their phonological structures. However, by applying acoustic analysis and proposing different measurements of vocalic and consonantal durations, Ramus *et al.* (1999) argued that the infants' differentiation between rhythms is not based on the phonological structures of languages, rather on rhythm correlates (vowels and consonants' durations). As such a new approach started to emerge aiming at quantifying the prosodic differences between languages.

This view that certain phonological features such as the syllable structure and language phonotactics have an influence on the rhythmic structure of that language was derived from Dauer's (1983) study of rhythm. Dauer's ideas were quantified by proposing indexes that enabled the researchers to apply phonetic measurements that may provide phonetic validation of that phonologically based view of rhythm. The rhythm indexes were applied in recent studies that investigated rhythm as a current approach that aims at measuring the

vocalic and consonantal durations in an attempt, as Grabe and Low (2002) put it, to quit seeking isochrony.

The Ramus *et al.* (1999) indexes included the total duration of vocalic intervals (%V), the standard deviation of vocalic durations (ΔV) and the standard deviation of consonantal durations (ΔC). According to these indexes the perfect isochronous indication is 0 for (ΔC) and (ΔV) and 50% for (%V). Ramus *et al.* (1999) presented three hypotheses for the study. The first was that the more complexities of the syllable structures of stress-timed languages would yield in more variations in the durations of the consonantal intervals. The second hypothesis was that vowel reductions yield in more variations in the vocalic durations of stressed-timed than syllable-timed languages. The last hypothesis assumed that the consonant clusters complexities together with the variable vocalic durations resulting from the patterns of stress would yield in a small percentage of the index in stressed-timed languages compared to syllable-timed languages. In order to test their hypotheses, they used their proposed indexes for measuring the vocalic and consonantal durations of eight languages that have always been categorized as belonging to a rhythm class, namely, English, Polish and Dutch as being stress-timed languages; Spanish, Italian and French as syllable-timed languages; and Japanese as a mora-timed language. The authors combined their results and presented them in three charts based on the values of each language as can be seen in figure 2.3:

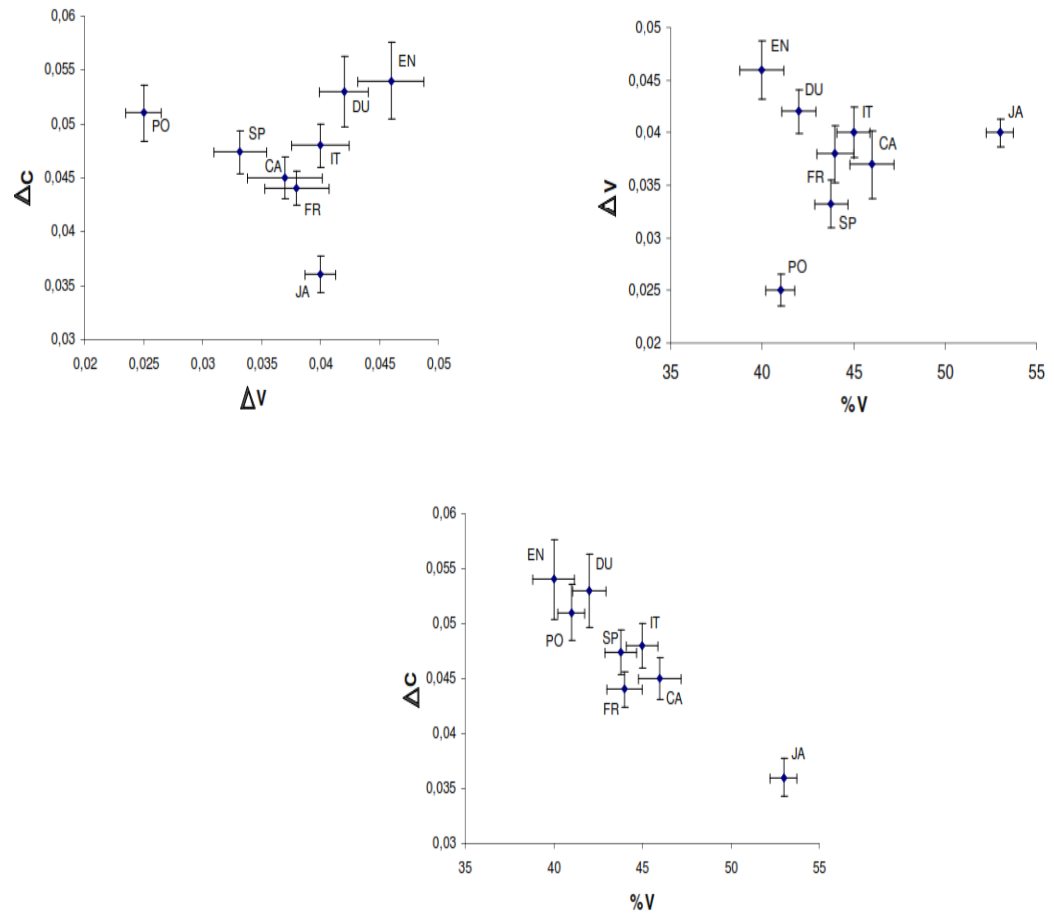


Figure 2.3: Distribution of languages over (ΔC), (ΔV) and $V\%$ scales (Ramus *et al.*, 1999, pp. 30-32)

The above figures show that the results coincided with their expectation in the sense that Dutch and English appear to be grouped together 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 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 values. As such, the ΔC and $\%V$ of their study displayed significant statistical differences while no differences were seen in ΔV owing to the low ΔV values which Polish exhibited in

comparison to the other syllable-timed languages, (Tan, 2011, p. 64). Ramus *et al.* suggested that combining ΔC and %V indexes can result in a more successful discrimination between stress timing and syllable timing features of languages from a controlled speech rate data, while ΔV seems to be very sensitive to certain factors. However, although their study provided basis for infants' linguistic and cognitive development, it did not take into account the influence of speech rate and it measures syllable structures rather than rhythm.

Prior to that, Low and Grabe (1995) compared Singapore English to British English. This study is not only important for comparing two varieties of the same language but it also introduced the Pairwise Variability Index (PVI). The PVI is used to measure the durations between successive vowels, including consonants, and other stretches of speech, (Ladefoged, 2006). It is a representative of vowel duration variability, where a lower PVI indicates more syllable-timed features of language while a higher PVI represents a more stressed-time language. These indexes involve measurements of the consecutive pair of vowels in an utterance. Then, the mean difference is calculated by adding the overall differences divided by the number of differences. The speech rate is controlled by having one less than the number of vowels. However, the PVI is not an indication of 'the absolute rhythm of a language' but it provides a possibility of comparing the rhythms of two or more languages, (Harris and Gries, 2011). Later on, Low, Grabe and Nolan (2000) used the index for further comparison between Singapore English and British English. It has been found based on the results of the PVI values that Singapore English appeared to be more syllable-timed than British English due to the slightly longer or less reduced vowels recognized in the former. Generally, both indexes became useful in acoustic studies, but whereas Ramus

et al. (1999) indexes focus on thorough measurements of utterances, Low *et al.*'s (PVI) is based on the vowel pairwise scale, that is to say, measuring the varied durations of sequential vowels rather than an inclusive variation of an utterance. According to Tan (2011, p. 62) the significance of both indexes of Ramus *et al.* and Low and Grabe is that they utilize certain features like, 'vowel reduction', 'syllable complexity' and 'stress-based lengthening' in measuring the rhythmic structures of different languages or different language varieties regardless of the intricate phonological combinations. However, in using Ramus *et al.* indexes, Low *et al.* (2000) found that ΔV can be compatible to PVI while %V is considered as been questionable due to the fact that differences between language varieties cannot be captured using this index. Low *et al.* also used a consonantal PVI similar to ΔC proposed by Ramus *et al.*, and it can be used for measuring languages that display both kinds of stress-timed and syllable-timed rhythmic structures. However they suggested that normalization should not be used with this index given that the differences in consonant durations are based on the sequences of speech sounds of a language. In comparing the PVI to the Ramus *et al.* (1999) indexes, Low *et al.* (2000) argued that the PVI is a better index for identifying the rhythm of languages. Besides, it appeared to be successful in capturing the rhythmic structures of languages like Polish and Catalan which have always been problematic for researchers.

In studying the rhythm of Singaporean English, Deterding (1994) used an equation called Variability Index (VI) to measure the duration of the whole syllable instead of measuring the duration of the vowel only, since according to him syllables might vary in length irrespective of vowel involvement. Besides, he studied the participants' conversational speech rather than their reading. His findings are similar to those of Low (1995) in that the

Singaporean rhythm can be classified as being more syllable-timed than British English given that they display more regular durations of successive syllables. He therefore suggested that it is plausible to think of these two varieties as being different in terms of stress-timing and syllable-timing scale. However, Thomas and Carter (2006, p. 339) consider this method as being problematic when using spectrograms given the difficulty in determining the divisions of syllables.

Grabe and Low (2002) applied the normalization of speech rate or $nPVI$ in calculating the vocalic durations of eighteen languages. They collected their data from one native speaker of each language reading an equivalent translation of the *North Wind and the Sun*. By having the x-axis representing the consonantal PVI and the y-axis representing the vocalic PVI, the so-called stress-timed and syllable-timed languages appeared to be separated in the graph while others stood in between. In other words, Thai, Dutch, German and English were classified as stress-timed languages, whereas Spanish, French, Tamil and Singapore English were classified as syllable-timed languages. Only Japanese appeared as a mora-timed language and the remaining nine languages were regarded as intermediate languages. Thus, Grabe and Low (2002) suggested that languages can be treated in a continual manner displaying more or less features of stress-timing or syllable-timing.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
	Th.	Du.	Ge.	BE	Ta.	Mal	SE	Gr.	We.	Ru.	Po.	Es.	Ca.	Fr.	Ja.	Lu.	Sp.	Ma n	
1 Thai							x	x	x	x	x	x	x	x	x	x	x	x	
2 Dutch							x	x	x	x	x	x	x	x	x	x	x	x	
3 German									x	x	x	x	x	x	x	x	x	x	
4 BE														x	x	x	x	x	
5 Tamil		o												x	x	x	x	x	
6 Malay														x		x	x	x	
7 SE																x	x	x	
8 Greek																x	x	x	
9 Welsh					o													x	x
10 Rumanian					o		o											x	x
11 Polish	o	o	o			o		o	o	o								x	x
12 Estonian	o	o		o	o	o	o	o			o							x	x
13 Catalan												o						x	x
14 French					o		o				o								
15 Japanese											o	o							x
16 Luxemb.											o	o							
17 Spanish											o	o							
18 Mandarin					o		o				o								

x = significant *n*PVI differences
o = significant *r*PVI differences

Figure 2.4: Differences between the languages based on *n*PVI and *r*PVI values (Grabe and Low, 2000, p. 8/16)

Figure 2.4 is adopted from Grabe and Low’s study and it shows, based on their findings, the non-dichotomous nature of language. It also displays that the number of the *n*PVI significant differences is double the number of the *r*PVI values which lead them to suggest that the *n*PVI seems to be more effective in discriminating between languages than the *r*PVI. However, although this study is quite significant for providing data from eighteen languages, the reliability of this study can be questioned for counting only on one representative of each language.

These studies came in favour of the assumption of discriminating 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.

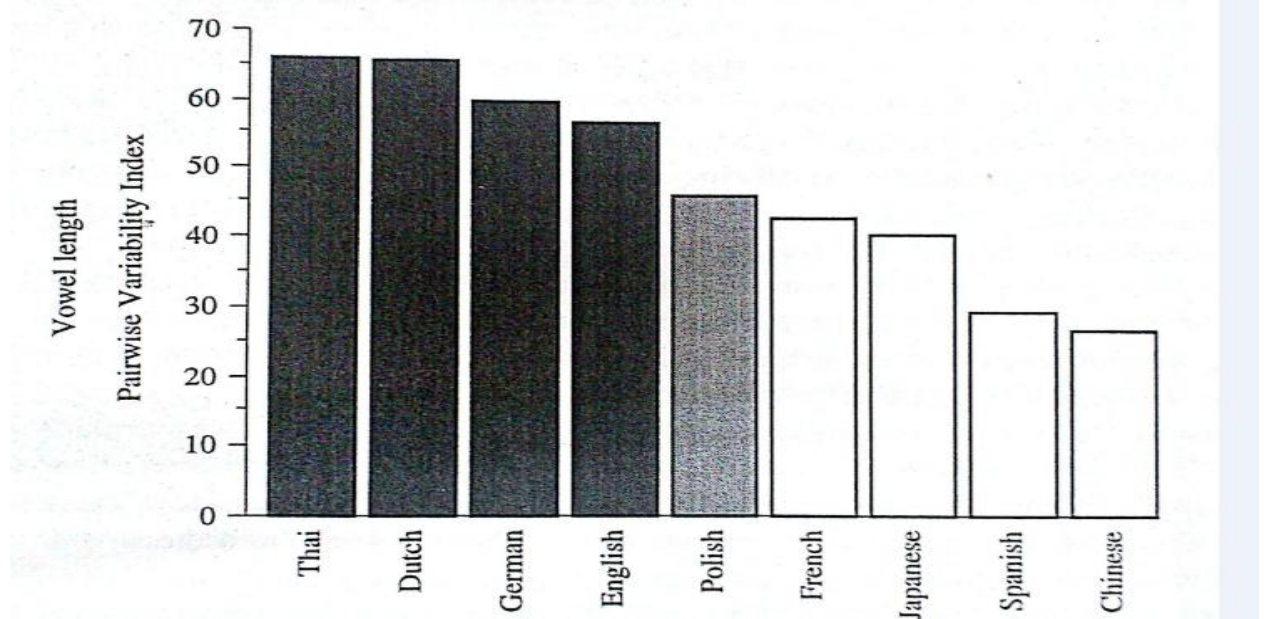


Figure 2.5: The PVI Values of Nine Languages, (Ladefoged, 2006, p. 246)

Figure 2.5 is an illustration of languages classification based on their rhythmic structures that has been analyzed using the PVI, which is adopted from Ladefoged (2006, p. 246). It shows variations from highly stress-timed languages represented by the dark columns to highly syllable-timed languages represented by the white columns. Yet, there is a grey column in between that represents Polish which displays mixed features of both stress-timed and syllable-timed rhythm. As such, these indexes are means of comparing between languages or dialects rather than merely classifying them, and they were used by researchers for this purpose.

Later studies that investigated rhythm attempted to modify or develop further indexes. In their comparison between the rhythms of two typologically different languages, namely, the stress-timed English and the syllable-timed Ibibio (a Nigerian language), Gibbon and Gut (2001) proposed the 'Rhythm Ratio'. They added a slight modification to the PVI by providing a 'more conventional range between 0 and 100', unlike the PVI which ranges between 0 and 200, and it was used for measuring the durations of both vowels and syllables, (Gibbon and Gut, 2001, p. 92). They calculated their ratios and added a binary classifier for a prediction of focal and non focal units of rhythm. Their findings showed evident differences between the stress-timing and syllable-timing features of both English and Ibibio.

Dellwo and Wagner (2003) investigated the rhythm of English, French and German by applying %V and ΔC to test the extent of their efficiency in terms of speech rate. Although speech rate seemed to affect the values of the indexes particularly ΔC , the grouping of stress-timed and syllable-timed languages in different places was recognized. As such, they suggested that in order to control speech rate, the ΔC should be divided by the average of the interval durations of consonants multiplied by 100, in what came to be known as Varco C which is influenced by the process of normalization found in the PVI. This index was examined by Dellwo (2006) providing a more evident distinction of English and German from French at different rates of speech concluding accordingly that this index is better than Ramus *et al.* ΔC in distinguishing between these languages.

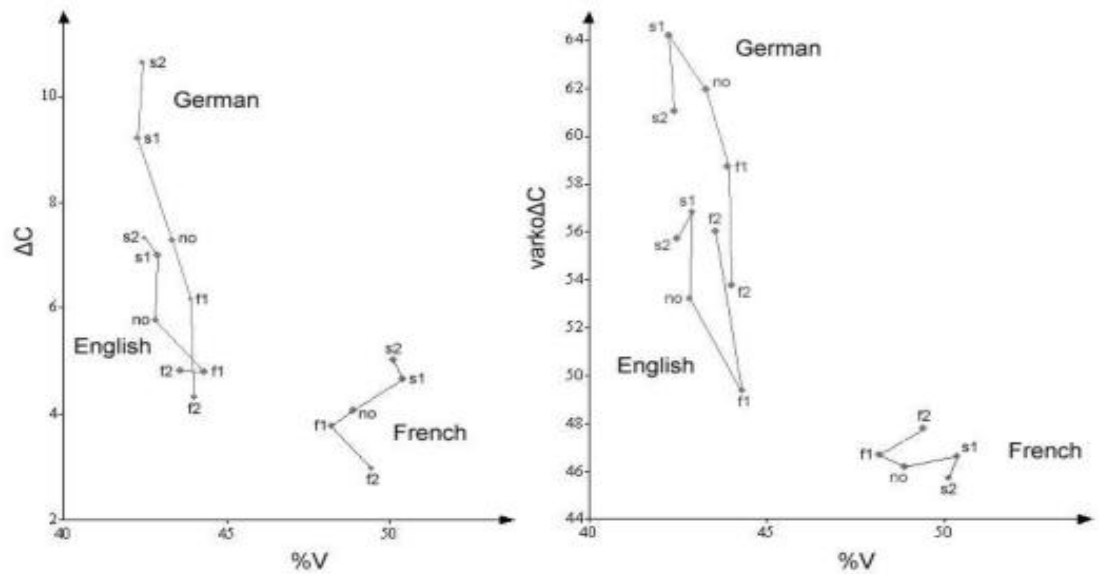


Figure 2.6: Discrimination between languages based on Varco C and ΔC scales (Dellwo, 2006, pp. 232-235)

Figure 2.6, adopted from the study of Dellwo (2006), shows that French appears more distinctly below English and German on the scale of Varco C than ΔC . Generally the distinction between the stress-timed and syllable-timed languages appeared to be more evident with Varco C which makes it more successful in discriminating between languages. Dellwo also used Varco V where the standard deviation of the vocalic durations is divided by the mean intervals of vowels multiplied by 100.

Influenced by the PVI, Wagner and Dellwo (2004) used an index called YARD (Yet Another Rhythm Determination) to measure the durations of syllables and found isochronous evidence at different prosodic levels of languages. Features of isochrony

appear in two or three syllables in English and Germany while they appear in three and four syllables in French and Italian.

Bertinetto and Bertini (2008) changed the *r*PVI into 'Control/Compensation Index' CCI in which the interval durations of each vowel and consonant is divided by the number of segments to differentiate between 'controlling' and 'compensating' languages which refer to languages like Japanese and English, respectively. The former displays controlled effort on each syllable, thus resulting in similar durations of segments. The latter on the other hand displays compensation for the prominent effort in stressed syllables with less effort in unstressed syllables resulting in more varied segmental durations.

In using different indexes to measure the rhythmic structures between and within languages, White and Mattys (2007a) strongly suggested that ΔV and ΔC need to be normalized owing to the relation between these indexes and speech rate, whereas %V did not display such a relation and thus could evidently distinguish between classes of rhythm. On the other hand, the normalization of Varco V and Varco C provide different results. Varco V could distinguish between and within classes of rhythm while Varco C did not display such distinctions which proves Low and Grabe right in that the *r*PVI should not be normalized. Thus, for measuring vocalic intervals, the normalized PVI was more effective while for consonantal intervals, the normalization seemed to be less reliable since it hardly showed distinction between rhythmic classes. Besides, the use of %V and Varco V collectively is quite complementary leading to more efficient and clear cut results. By applying different acoustic indexes to capture the rhythmic differences between languages,

Wiget, White, Schuppler, Grenon, Rauch and Mattys (2010) tried to study the effect of speakers, sentence materials and measurers as being three major sources of variation in scores of rhythm indexes. It has been shown that sentence materials are considered as the most important and largest source of variation, while speakers are seen as a moderate source of variation. However, variations of indexes scores caused by measurers are much smaller than those caused by sentences and speakers. This comes in line with the findings of Loukina, Kochanski, Shih, Keane and Watson (2009) who applied fifteen indexes in their study with calculating each of them in three different ways, namely, using calculations for every sentence then taking the average of part of the results; using calculations for every sentence, without including the segments that occur before pauses to avoid lengthened utterances and then taking the average of part of the results; and finally, using calculations for the whole text. As such, forty five indexes were applied in a process of computed segmentation to provide algorithmically a perfect prediction of a language by combining indexes. They found that there are fundamental differences in languages which make it impossible to draw conclusions based on the rhythm of a material that does not display clear segmental information. Besides, although some indexes can successfully achieve a classification or rather distinguish between languages such as the combination of V% and *n*PVI, the effectiveness of indexes are highly relevant to the languages compared in one study, accordingly, different groups of languages lead to different results which need to be considered in the comparison of the findings.

Schmid (2001) applied the Ramus *et al.* indexes to investigate the rhythmic structures of Swiss German, German and English, which are categorized with stress-timed languages, and Spanish, Italian and French, which are categorized with syllable-timed languages. The

findings emphasized the reliability of ΔC and ΔV for measuring rhythm. However, this was not the case with %V which provided results that contradicted the author's expectations given that it displayed high values for English and low values for French. Benton, Dockendorf, Jin, Liu and Edmonson (2007) compared between English and Mandarin using both the Ramus *et al.* indexes and the PVI. Both indexes classified the former as a stress-timed language and the latter as a syllable-timed language. Tortel and Hirst (2008) investigated the rhythmic structures of English spoken by French by applying the PVI. The participants were divided into four groups, with the first consisting of French speakers with the least abilities of English speaking skills, or in other words non-English students, the second group consisted of English students (freshmen), while seniors were included in the third group. The fourth group consisted of English native speakers to compare their values with those of the French speakers. It was found that the *n*PVI was more successful in reflecting the speakers' abilities in English where the values seemed to raise evidently from the first to the last group. The *r*PVI, on the other hand, did not display such consistency. Bunta and Ingram (2007) also used the PVI to investigate the rhythmic structures of children who are bilingual speakers of Spanish and English. They found that the younger the participants, the greater the variations in the rhythmic structure occur between monolingual and bilingual speakers. However, these variations seemed to decrease among older speakers.

Deterding (2011) used the PVI to measure the rhythm of Malay compared to Southern British English using the terms stress/syllable based rhythm, and found that although Malay is more syllable-based than British English, there might be some problematic issues regarding this result such as the speaking tempo and syllabification which is highly prone to

subjectivity. It was shown that Malays take longer time in reading a given material, however, owing to greater number of syllables in Malay, this can lead us to assume that the speakers produced syllables in faster tempo per second. But such an assumption can be misleading given that Malay syllable structure does not allow consonant clusters and therefore the syllables are less complicated than those of English. As for syllabification, the Malay word 'kuat' and the English word 'traveller' from the '*North Wind and the Sun*' text can lead to different PVI values if the former is treated as monosyllabic or disyllabic and the latter as disyllabic or a trisyllable word. As such, despite the effectiveness of PVI, caution is quite required in drawing conclusions about rhythm especially in terms of tempo and subjectivity of syllabification.

Asu and Nolan (2006) measured the durations of vowels and consonants in Estonian compared to English by applying the PVI. Although Estonian has already been classified by Grabe and Low (2002) as a syllable-timed language, this study presented significant results based on syllable and foot PVIs. Despite the traditional assumption that a stress-timed language like English controls for foot rather than syllables, it appeared in their study that English displayed more variations at the level of syllables than Estonian, while less variations were seen at the level of foot which were comparable to the few variations seen in Estonian making the latter controls for both syllable and foot. They wrapped up their argument by suggesting that the PVI measurements at the level of syllable and foot are separate from those at the level of stress and thus a combination of both measurements can lead to a more reliable analysis of rhythm.

Mok and Dellwo (2008) studied the speech rhythm in Cantonese and Mandarin as well as Cantonese-accented English and Mandarin-accented English, using a number of indexes including the PVIs, Varco C and Varco V and the Ramus *et al.* indexes. Their findings revealed that Cantonese show higher values than Mandarin given that word stress does not occur in the former. The results also show that both accents are influenced by the syllable-timed features of their mother tongue. The participants did not display any reduction of unstressed syllables as native speakers of English do, yet, their vowel lengthening may not be attributed to the influence of Mandarin or Cantonese, but it can also be a result of a 'slower speaking rate' due to the participant's lack of fluency in speaking English. However, although these indexes were successful in identifying the rhythmic structures of Cantonese and Mandarin, the identification of the rhythm of the two foreign accents of English appeared to be problematic due to the fact that both of them are perceptually syllable-timed, yet, appeared as stress-timed by a number of rhythmic indexes. These findings come in line with the results of Lin and Wang (2008) who studied the Mandarin accent of English and found that it showed an *n*PVI value of (60.15), which is higher than Grabe and Low's (2002) values of both English (57.2) and Mandarin (47.61). This might imply that their participants have strong command of English, probably because the data were elicited from Chinese participants in Canada, however, the results revealed that the speakers were not quite competent in consonantal durations of English due to the complexity of English syllable structures that allow three and four consonant clusters unlike the Mandarin syllable structure that allows only for one consonant in syllable onset and coda. Several other studies have also employed different indexes to compare between foreign accents of English and different varieties of languages finding that there are noticeable cross-dialectal variations in rhythm, such as those found by Low *et al.* (2000)

between the two English varieties, British English and Singapore English and those found by Jian (2004) between American English and Taiwan English. White and Mattys (2007) also found that Welsh and Orkney English display significant rhythmic variations compared to the rhythmic structure of the Standard Southern British English. Barry, Andreeva, Russo, Dimitrova and Kostadinova (2003) found that different varieties of Italian, such as Naples, Pisa and Bari, display different types of rhythm. Similar differences were found by Frota and Vigario (2001) in comparing Portuguese spoken in Europe and that spoken in Brazil. Ghazali *et al.* (2002) also found some rhythmic variations between the eastern and western dialects of Arabic.

2.6 PVI as an Analytical Framework of the Study

Different rhythmic indexes have been proposed by researchers to measure the durational variations between phonetic or phonological elements in speech. Some of these indexes measure the global variations in utterances, such as the Ramus *et al.* deltas which measure the standard deviation of durational intervals between consonants or vowels, or Varco C and Varco V which divide the standard deviation by the mean durational intervals of consonants and vowels and multiply them by 100. The PVI on the other hand measures the sequential durational variations between adjacent pairs of vowels or consonants and calculates the mean of those pairwise variations. It differs from the standard deviation in that the latter measures all the occurring pairs in an utterance, whether they are successive or not; the PVI, however, measures the durational variations between successive pairs of vowels or consonants in an utterance. The difference between the two indexes can be illustrated in figure 2.7 which is adopted from Mairano 2010, p. 40):

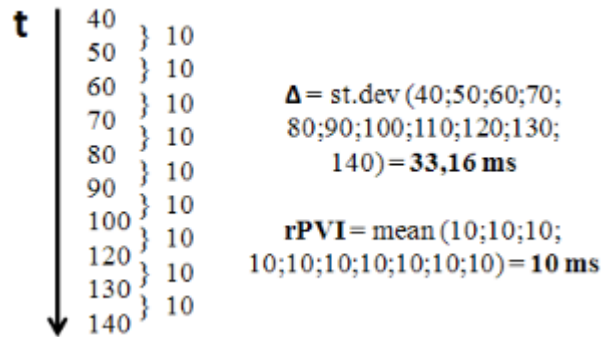


Figure 2.7: Difference between the Standard Deviation and the PVI (Mairano 2010, p. 40)

Figure 2.7 shows a presumptive durational sequence of vowels ranging from a higher to a lower tempo. A rise of 10 ms in the tempo would result in different values of the PVI and the standard deviation. The latter measures all the pairs while the former measures only the sequential pairs variations (which, in this hypothetical succession is constantly 10).

The PVI involves the *n*PVI which is used in cases that involve sensitiveness to speech rate such as the vocalic durational intervals, and is normalized in order to control speech rate, by dividing it by the sum of the intervals. As for the *r*PVI, it is used in cases where there is no sensitiveness to speech rate such as the consonantal durational intervals. Grabe and Low (2002) suggested that normalization of the *r*PVI should be avoided in measuring the consonants given that the size and differences of the consonantal intervals are influenced by the phonotactics of the language and thus, normalization might result in eliminating the rhythmic differences. The application of the PVI for measuring the acoustic correlates of rhythm would result in high values if the language under investigation employs stress

timing features. However, if a particular language employs features of syllable timed rhythm, the PVI values will be low.

The PVI is applied in this study to measure the vocalic and consonantal durational intervals of YAD and JAD. Many studies found that the PVI is a better index for measuring rhythm due to its reliability and robustness in discriminating between languages or even between language varieties. Grabe, Post and Watson (1999), for instance, applied this index in their study and found that it was quite efficient and successful in distinguishing between the stress timing features of English from the French syllable timing features. Similarly, Low *et al.* (2000) used the PVI in comparing between British English and Singapore English. The results seemed to conform to the author's predictions in displaying significant differences between both varieties. Comparison between these two varieties of English became an object of study by Grabe and Low (2002), but this time by adding another dimension to the PVI which is the normalized version of the index or *nPVI*. By applying this index in measuring the durational vocalic intervals of eighteen languages, the authors were able to provide a continuum that ranges between the highly stress-timed Thai to the opposite syllable-timed Chinese. A number of languages appeared in between for displaying intermediate features of rhythm, thus conforming to Dauer's (1987) remarks about rhythm. Loukina *et al.* pointed out that whenever a comparison between different rhythmic indexes is involved, the PVI appears to be one of the best and most efficient indexes in achieving high success rate in discriminating between languages (2009, p. 1533). White and Mattys (2007) compared between the stress timing features of English and Dutch and the syllable timing features of French and Spanish by applying a number of indexes in analyzing their

data. They found that the *n*PVI and Varco V appeared to be the most effective indexes in their discrimination between and within the categories of rhythm.

2.7 Speech Rhythm in Arabic

Earlier studies of the rhythm of Arabic described the language as being characterized by an isochronic nature where prominent syllables tend to appear at fairly regular intervals no matter how many unstressed syllables appear between the stressed ones. However, the previous studies of Arabic rhythm were mainly based on auditory impressions and the language was categorized as being stressed timed accordingly. In his study on the perception of rhythm, Miller (1984) asked hearers from different backgrounds to categorize languages according to the structure of their rhythms and Arabic was categorized with stress-timed languages. Different other studies started to investigate the rhythm of different Arabic dialects. Hassan (1955, p. 163) described the rhythm of Egyptian Arabic dialect as being stress-timed since the range between the stresses of the language appears to be roughly the same. Soraya (1966, p. 23) also classified Egyptian dialect with stress-timed languages based on the following examples that describe the rhythmic features in Egyptian dialect:

gad| kallem a| khuwh (Gad spoke with his brother)

samy| kallem a| khuwh (Samy spoke with his brother)

shalaby| kallem a| khuwh (Shalaby spoke with his brother)

samy bey| kallem a| khuwh (Samy is speaking with his brother)

Soraya (1966, p. 23)

Soraya argued that, in spoken speech, the above utterances can be divided into three durationally equal segments based on their stresses, regardless of the number of syllables that appear between the sequences of stress.

Rammuny (1966, p. 31) also suggested that the duration of each group of stresses in Jordanian Arabic dialect is almost equal, thus it is regarded as stress-timed. In dealing with the rhythm of Palestinian dialect, (El-Isa, 1982, pp. 91-2), described it in terms of the folk dance performances where the dancers display evident regularity in stomping their feet in a manner that presumably matches with the stressed syllables that occur in the words of the song. However, it needs to be noted here that in reality, poetry and songs do take liberties with the structure of a language.

Heliel (1977), however, argued that, unlike the other Arabic dialects, those in North Africa like Moroccan and Algerian are syllable-timed and it is this feature that makes these dialects unintelligible to the other Middle Eastern Arabic dialects.

These studies were mainly based on perceptual findings on the basis of isochrony postulating that stress is the most prominent feature in Arabic that occurs at regular intervals. However, after the advent of a new approach to rhythm which involves acoustic

analysis of a language and application of indexes that measure the vocalic and consonantal durational intervals, researchers started to move towards acoustic investigations of Arabic rhythm to find out whether their findings will conform to those provided from the previous perceptual analyses of the language. In comparing between Arabic, Japanese and English rhythms, Tajima *et al.* (1999) found that speakers of Arabic and English show similar patterns of rhythmic modes which are different from those of Japanese speakers. Arabic and English speakers appeared to be quite careful with the stressed syllables, producing them at 'simple harmonic phase', while in the case of the Japanese speakers, it was the final syllable of the phrase that was placed at simple harmonic phases, (pp. 286-7) However, although both English and Arabic are said to be stress-timed, they seem to exhibit systematic differences in the sense that Arabic shows a higher level of deviation from the strict rhythm isochrony, making English more stress-timed than Arabic. However, the Arabic speakers in this study are represented by Jordanian speakers only, which makes it difficult to make a generalization about Arabic language which has a large number of other dialects.

Based on the findings of the perceptual experiments conducted by a number of researchers, listeners with no knowledge of Arabic can successfully distinguish between the dialects of North Africa and those in the Middle East. Accordingly, it is most likely that the patterns of rhythmic structures vary between different Arabic dialects. Generally, Arabic scholars divide all the dialects of Arabic into either eastern Arabic dialects (represented by the dialects of the Middle East) or western Arabic dialects (represented by the dialects of North Africa), with describing the latter as being more stress-timed.

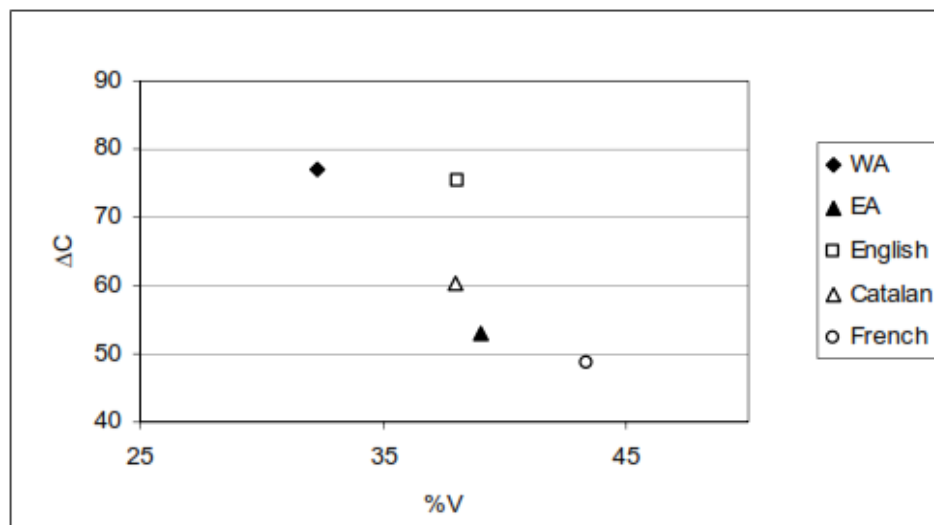


Figure 2.8: Comparison of the eastern and western Arabic dialects with other Languages (Hamdi *et al.*, 2004, p. 2/4)

Figure 2.8 demonstrates a comparison of the %V and ΔC of the eastern and western Arabic dialects with other investigated languages like English, French and Catalan. It appears that the western dialects display evident stress timing features comparable to those of English.

In their study of different Arabic dialects, Hamdi *et al.*(2004) applied the %V, ΔC and the PVI to compare between the Arabic dialects in order to determine whether Arabic rhythm has a separate category based on the type of its syllables or it forms a continuum with more or less features of syllable timing and stress timing. They revealed that eastern and western Arabic dialects do display different rhythmic structures, and although all Arabic dialects can be categorized with stress-timed languages, some dialects are considered to be more stress-timed than others. Moroccan and Algerian Arabic dialects show frequent vowel reduction, difference in vowel duration and syllable complexities, similar to the typical stress-timed language ‘English’. Jordanian, Syrian and Egyptian dialects, on the other hand,

display less features of stress-timed rhythm. Tunisian Arabic dialect, however, stands midway in the eastern-western Arabic dialect continuum, where it displays features similar to both classes of dialects. Tunisian vowels are longer than those in Moroccan and Algerian dialects, but slightly shorter than the eastern Arabic dialects vowels. It also displays less syllable complexity than the other western Arabic dialects. Thus, Hamdi *et al.* (2004) suggested that rhythm classes may not always be polar opposites, but they can also overlap in some dialects or languages.

In applying the Ramus *et al.* indexes for studying the rhythm of the different Arabic dialects in Morocco, Algeria, Tunisia, Egypt, Jordan and Syria, Ghazali *et al.* (2007) found that the %V of dialects gradually increases as we move from west to east while on the contrary, ΔC decreases. However, dialects of Tunisia and Egypt which are located in the intermediate zone exhibit intermediate values, but it is shown that with regard to ΔC , Tunisian and Egyptian Arabic dialects are closer to Middle Eastern dialects than those in North Africa. Using a *t*-test for comparing between dialects, they found that there is no significant difference between Jordanian and Syrian Arabic dialects which both belong to the same region. However, significant differences occur between dialects which are located in different extremes of the Arabic-dialect continuum such as Syrian and Moroccan which show a probability value of (0.001). As such, the degree of significance is fairly relevant to the distance between the dialects under investigation. It is worth mentioning that Tunisian and Egyptian dialects, which are almost at the centre of the continuum display significant differences only when compared to the Moroccan dialect. The following charts show the distribution of the Arabic dialects in two and three groups on the scale of %V and ΔC based on the study of Ghazali *et al.* (2002).

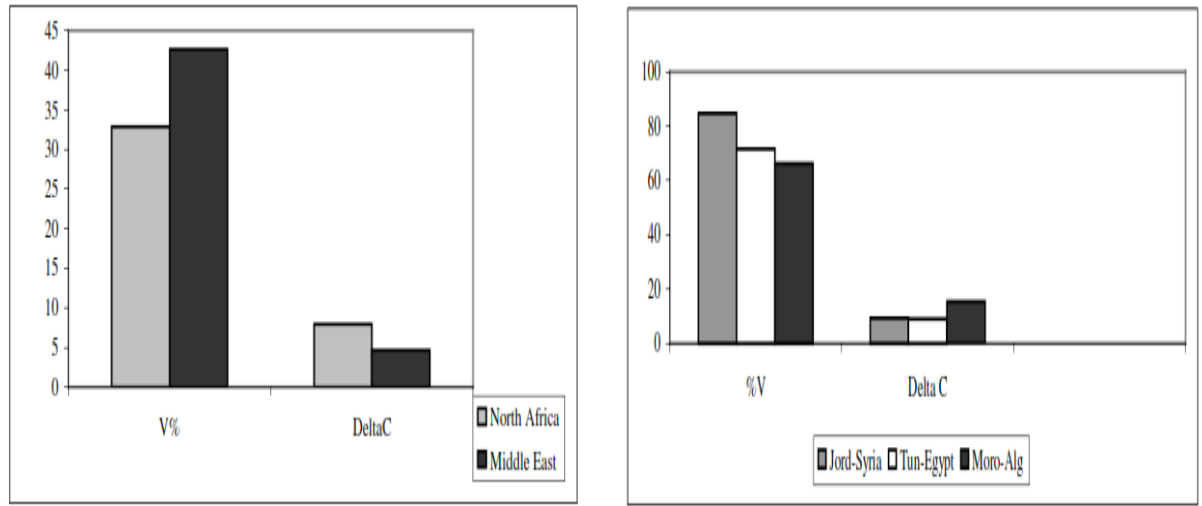


Figure 2.9: Distribution of Arabic dialects on %V and ΔC scales (Ghazali *et al.*, 2002, p. 333)

In his study of the Tunisian Arabic rhythm, Barkat (2000) found that this dialect has a %V comparable to that of the North African Arabic dialects but a ΔC similar to the Middle Eastern Arabic dialects. They produce vowels which are somewhat longer than the Algerian and Moroccan vowels, but are markedly more reduced than the vowels of Syrian and Jordanian dialects. However, they do not have the same complex syllable structures of the other North African Arabic dialects.

2.8 Conclusion

This chapter presented a description of Arabic language by explaining the syllable structure as well as the vocalic and consonantal systems of SA, YAD and JAD. The concept of

prosody and some of the prosodic or suprasegmental features of human speech are also introduced in this chapter. Different definitions of rhythm have been presented and different approaches to rhythm have been discussed. The chapter also presented the current approach to rhythm that is associated with different rhythmic indexes by referring to studies that have investigated rhythm by applying those indexes. Studies that investigated speech rhythm in Arabic have also been reviewed. The theoretical and analytical frameworks were also presented in this chapter. The following chapter will present the methodology of the study by providing a detailed description of the research design, materials, participants and methods of data collection and data analysis.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The previous chapter presented a description of Arabic language and its dialects and proceeded with a review of prosody, rhythm and rhythmic index. This chapter provides a description of the research design and the analytical framework which is used for the analysis of the data. Descriptions of the materials used in the study as well as the background of the participants are also given. The methods of data collection and analysis are provided with illustrations that clearly display the process and consistency of the segmentation of the vocalic and consonantal intervals.

3.2 Research Design

The present study attempts to conduct an acoustic investigation of the YAD rhythm and tries to identify its position among the other investigated and classified Arabic dialects. As mentioned earlier, researchers have divided Arabic dialects into eastern Arabic dialects represented by some dialects of the Middle East, namely, Jordanian, Syrian, Lebanese and Egyptian, and western Arabic dialects represented by some dialects of North Africa, namely, Moroccan, Algerian and Tunisian. These two categories of Arabic dialects display features of stress-timed rhythm but with some disparities in the level of stress-timing features. Therefore, the Arabic dialects are treated in the form of a continuum with the

western dialects appearing in one extreme of the continuum displaying more stress-timing features than the eastern dialects in the opposite extreme or the intermediate dialects that carry features from both categories. Being one of the Arabic dialects that has not yet, to the best of the researchers knowledge, been identified in terms of speech rhythm, YAD is investigated in the present study aiming at identifying the rhythmic features of this dialect. To achieve this goal, the study investigates the acoustic correlates of YAD rhythm and compares them to those of JAD which is classified as an eastern dialect. For this purpose, Yemeni and Jordanian native speakers were recorded individually and their output was measured and analyzed using the Pairwise Variability Index (PVI). According to researchers like Grabe *et al.* (1999), Low *et al.* (2000), Grabe and Low (2002), White and Mattys (2007) and Loukina *et al.* (2009), it is one of the most successful indexes and has shown to be quite effective as far as measuring rhythm is concerned. It is useful for discriminating between languages and controlling the problems that are brought about by differences in speech rate. The PVI is a calculation of the durational differences between successive pairs of units, namely, the durations of vocalic or consonantal intervals, followed by adding those pairwise differences and taking the mean difference between those pairs. As such, the greater the difference in the vocalic or consonantal durations, the higher the PVI values will appear, indicating stress-timed features of a language. On the contrary, the lesser the difference, the lower PVI values will be, denoting a syllable-timed language. These differences are influenced by certain factors such as differences vowel duration or syllable complexities.

Following the previous studies of Arabic dialects rhythm which measured both the consonants and vowels of the dialects in question, the current study also measures these

two parameters using the normalized version of the PVI (*nPVI*) to measure the vocalic durations and the raw PVI (*rPVI*) to measure the consonantal durations. The purpose of using a normalized PVI for measuring the durations of vocalic intervals is that vowels by nature are prone to changes depending on the speech rate, therefore, a normalized version of the PVI is required. On the contrary, Grabe and Low (2002) suggested that normalizing speech rate is inadvisable in the case of consonants since they are not sensitive to speech rate and that the changes that occur in the duration of consonantal intervals is a result of the language phonotactics and not the variations of speech rate. Given that numbers are used to describe the language, the study can be said to be descriptive in nature.

3.3 Materials

In order to fulfill the objectives of the research, the study makes use of a reading passage as well as spontaneous speech. As far as spontaneous speech is concerned, a number of topics were proposed for the study to get the participants to choose one which is of most interest to them. The choice of the topics was mainly influenced by the gender variable in the study and indeed they seemed to match the interests of the participants. With respect to the reading passage, the current study makes use of a SA translation of the well-known English fable '*The North Wind and the Sun*' adopted from Thelwall and Sa'addedin (2003, p. 54). This passage was used by Grabe and Low (2002) and shown to be useful for phonetic work in English as well as in other languages translations, thus it was employed in this study as the material for the participants to read from and record. According to Deterding (2006), this passage works well for measuring the duration of vowels and consonants, and several data have been recorded and analyzed using this text. However, Deterding (2006) suggested that the English version of '*The North Wind and the Sun*' has some shortcomings in that it

consists of words that might be perceived differently in terms of the number of syllables such as the word ‘traveler’ which occurs four times in the passage. Tan and Low (2013, p. 7) also pointed out that this passage can be problematic as it has several initial approximants or words that have a sound /r/ in a syllable making it difficult to get accurate measurements of the vowel durations. As for the Arabic version of the passage, the most noticeable feature was the several occurrences of glottal stops followed immediately by a short vowel as noticed in words like /ʔayin/, /ʔaqwa/, /ʔuxra/, /ʔið/, /ʕabaʔatin/, /ʔidzbari/, /ʔaqsa/, /ʔila/, /ʔan/, /ʔusqita/ and /difʔiha/. Although we find in the passage several occurrences of the glottal stops which are followed by vowels, there is no single instance found where the glottal stop is followed by a consonant or appears in the syllable coda despite the fact that these cases occur quite frequently in Arabic language in such words as /beʔr/ (well), /jeʔt/ (you want), /luʔluʔ/ (pearls) /qaraʔ/ (read) and /defʔ/ (warmth). Had such occurrences appeared in the passage, a better observance and more accurate measurement of the glottal stop duration would have been achieved.

3.4 Participants

Ten participants were selected for this study, all of whom are postgraduate students in Kuala Lumpur, Malaysia. Five are Yemeni speakers of Arabic, who came from Ta'iz city in Yemen; and the other five are Jordanian speakers of Arabic, who came from Amman city in Jordan. The ten participants were females to keep the gender variable consistent, and their ages ranged between 25-35. According to Tannen (1990), female speakers always exhibit eagerness to talk and their goal in language use is to constantly gain intimacy unlike men who often tend to be silent and use language mainly to preserve their independence.

Counting on these motivations, the researcher expected to generate sufficient data from the female participants and find more cooperation from them.

In order to fulfill the objectives of the study, the participants needed to be educated primarily in their own countries and had lived in Malaysia for less than one year to avoid as far as possible the potentiality of linguistic and extralinguistic influences over their output. However, although the researcher initially aimed at finding participants from the same educational institution, in practice, this was not easily applicable since it was difficult to find sufficient Yemeni and Jordanian speakers who fulfill the conditions of the study. As such, the participants were selected from different universities in Malaysia. The Institute of Postgraduate Studies, (IPS) in University of Malaya (UM) provided the names of all the Yemeni and Jordanian students in the university. This was followed by a process of contacting these students to identify the suitable participants. Given the insufficient number of the required participants from one university, the option was extended to participants from other universities such as University Kebangsaan Malaysia (UKM) and International Islamic University Malaysia (IIUM). Generally, the homogeneity of the speakers in terms of gender, age and educational background was intentional in order to obtain a comparable set of data from speakers of both dialects. Barry, Andreeva, Russo, Dimitrova and Kostadinova (2003) and Arvaniti (2012) suggested that different and diversified speaking styles can be observed from speakers of different backgrounds, positions and occupations in society such as students of different levels and degrees, teachers, professors, artists and others. Therefore, this study targeted homogeneous speakers to avoid diverse and different kinds of speech output. Although different styles of speech might be useful to reflect the

diversified nature of language, however, in this study, the work does not involve a large number of speakers and such differences within a small group of five speakers of each dialect might be problematic in the findings.

3.5 Data Collection

For data collection, the study adopted two tasks of those employed by Labov (1966). The Labovian model aims at describing speaker's production in different situations and focusing on the speaker's diversified styles that result from his mental and psychological influences (Alias, 2003, p.115). William Labov (1966) employed four tasks:

1. Passage reading
2. Word list reading
3. Spontaneous speech
4. Minimal pairs reading

Two of these four tasks were chosen for this study, namely, 'passage reading' and 'spontaneous speech'. Presenting the participants with passage reading is an important way of collecting data for a work oriented toward Phonetics or Phonology, given the fact that passage reading will give the researcher a chance to record the same sounds and sentences for every individual speaker, (Feagin, 2008, 31). However, one task was not sufficient to fulfill the purpose of the study. Since the research is about an investigation of a dialect which is more spoken than being read, spontaneous speech recording was quite essential for the study to elicit useful data from a sort of natural speech output. Besides, it has been argued by Arvaniti (2009) that data elicited from a reading text may not represent the rhythm of a particular language since the values of spontaneous speech are different from

those of reading speech due to the different prosodic characteristics employed in reading and speaking. However, Tan and Low (2013, p. 7) argue that although spontaneous speech seems to be quite suitable for measuring rhythm, a fixed text that gets the participants to read the same material might be more useful for the purpose of comparing between dialects as long as the speech output does not seem to be quite artificial and unnatural. Therefore, this study utilized both tasks to compare between spoken and read prosody to observe data from both spontaneous and self-conscious speech.

Alias (2003) employed the Labovian model to investigate the phonological variation among Malaysian learners of English. Likewise, Al-Shuaibi (2009) adopted the same method for collecting his data to study the English phonotactics of Yemeni speakers of English. And in all phonetic work, we find that all the above mentioned tasks are essential to acquire data that can be analyzed both perceptually and acoustically. Researchers like Ramus *et al.* (1999), Low *et al.* (2000), Deterding (2006), Grabe and Low (2002), Ghazali *et al.* (2007), White and Mattys (2007) and Loukina *et al.* (2009), measured utterances from read texts. In this study, both passage reading and spontaneous speech were used for obtaining stream of information or continuous speech that is required for investigating rhythm.

3.5.1 Passage Reading

In order to fulfill the objectives of the research, the speakers were asked to speak into a microphone and audio recorded directly on a laptop in a quiet room, away from any kind of noise and distraction. Each participant was recorded individually and was presented with a written passage, namely, a SA translation of the well-known English fable ‘*The North Wind*

and the Sun'. The participants were allowed to take the time they needed to read the passage before recording and were recorded only when they felt comfortable about it. Since the vowel sounds appear in Arabic in the form of diacritics up or below the consonant sounds according to the case of the word in which it appears, these diacritics were carefully marked in the passage to assure that all the /a/, /i/ and /u/ sounds are produced by the speakers without any substitution of one vowel sound with another. Thus, the participants were given the instructions to read the passage as it is without adding, deleting or substituting any of the diacritics to avoid different articulations of the same vowel sound. For this reason, the participants were asked to read the passage over and over again to get it recorded properly, and if a participant would produce a wrong vowel sound, she would kindly be requested to reread the whole passage until no mistakes were made. As for consonants, they were much less problematic since they are not subject to changes no matter where they appear in a word or a sentence. However, one instruction was given to the participants regarding consonant sounds. It was related to the sound /t/ in the word '*tatagadalu*' (disputing) of the first sentence, where it was noticed that some speakers tended to delete the second /t/ sound due to the very short vowel sound that stands between both the /t/ sounds, thus producing it as '*tagadalu*'. Therefore, the participants were asked to pay much more attention when they came to produce this word.

3.5.2 Spontaneous Speech

In the second stage, the participants were also recorded individually, but this time using their spontaneous speech which was a more challenging task than the previous one. Prior to recordings, all the instructions were given by the researcher to the participants accompanied

by spontaneous conversations in their mother tongue. The purpose is to create a comfortable situation for the participants and reduce any kind of stress or tension that might affect their articulations. They were asked to have a full idea and complete thought of what they were going to say in order to avoid unnecessary pauses and linguistic stumbles.

The participants were asked to talk for at least ten minutes about either a memorable vacation, marriage, studying, working or cooking. According to Milroy and Gordon (2003, p. 60-61), the topics that a researcher chooses should be of interest to the participants, and since peoples' concerns are different, they should be presented with a variety of topics to choose from. Indeed, the participants were asked to select one of the aforementioned topics which were chosen carefully to match the general feminine taste and it was expected that any female participant would be interested in talking about at least one of these topics. The participants seemed to be interested in the topics given to them, however, the problem was in that they found it awkward to be recorded talking using their own dialect for ten minutes which was, according to them, a much more difficult and embarrassing experience than passage reading. That is why the participants were given all the time they needed to get themselves ready and to think and recall their experiences before speaking and recording, and if necessary, the participants were allowed to get themselves recorded alone in the room without even the presence of the researcher herself, as happened with one of the Yemeni participants who preferred to speak and record herself alone in the room.

3.6 Data Analysis

All the recorded data was saved in WAV form and then divided into utterances for each participant's reading text and spontaneous speech output. The utterances were identified by expressing a single intonational phrase, that is to say, the selection of the utterances was mainly based on identifying the positions where a pause was inserted. With regard to the read text, the passage was divided into eight utterances with some disparities in the number of items in each utterance. As for the spontaneous speech, eight utterances were also chosen to have a matching set or similar number of utterances as those of the reading text. Similar to the method used for the identification of utterances for the read text, identification of utterances was done through the pause locus. And as recommended by Deterding (2001) and Tan and Low (2013), utterances that included stuttering or unnecessarily lengthened vowels, or even some times consonants, were avoided. For instance, there were some cases where one of the participants tended to lengthen some vowels and consonants which appeared to be, as noticed by the researcher, to be a stylistic characteristic of the speaker to stress on a particular idea. Figures 3.1 and 3.2 show how the consonant sound /l/ and the short vowel sound /u/ were unnaturally lengthened by the speaker.

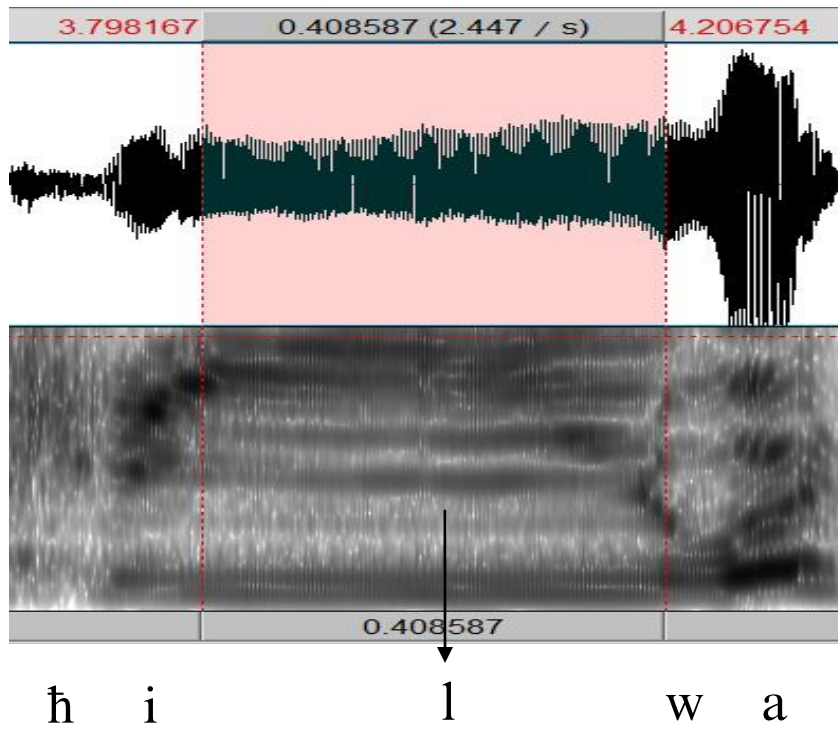


Figure 3.1: Spectrogram of the production of the sound /l/ by Yemeni speaker 4

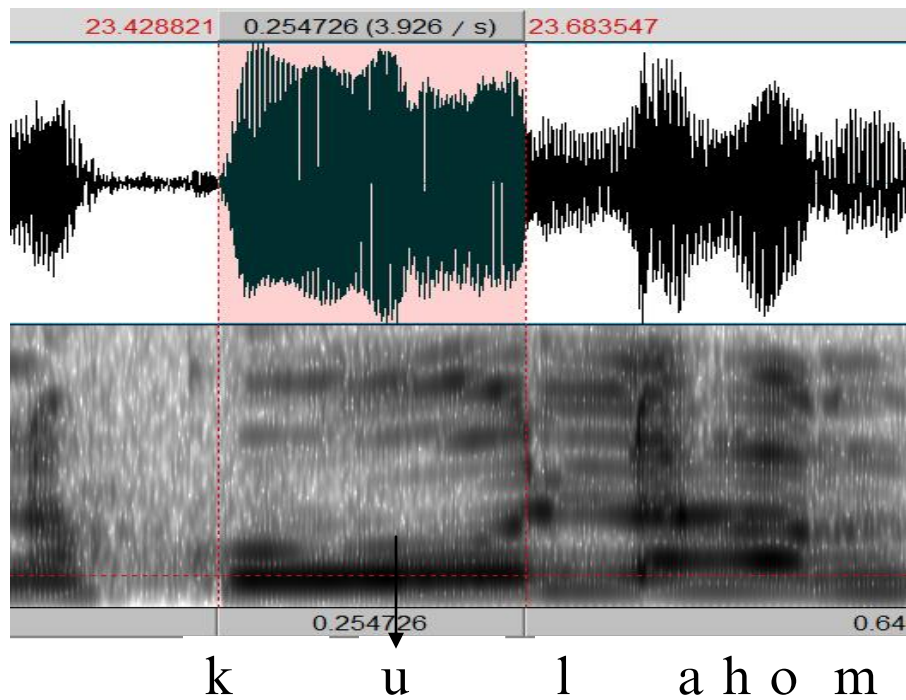


Figure 3.2: Spectrogram of the production of the sound /u/ by Yemeni speaker 4

In figure 3.1 the participant lengthened the sound /l/ which appeared in the word *hilwa* ‘beautiful’ when she was talking about her memorable vacation. The lengthened /l/ sound was used by the speaker to emphasize on how beautiful the vacation was. In the same vein, figure 3.2 shows that the short vowel sound /u/ was unnecessarily lengthened in the word *kulahom* ‘all of them’ to underscore the idea of the involvement of all her family in enjoying that particular vacation. As a result, the vowel sound /u/ was realized as /u:/ due to the sound lengthening tendencies of that speaker. Some cases of long pauses did also occur when speakers could not find suitable words to express themselves, thus they tended to take these pauses to gather ideas or even end the sentence abruptly. Fortunately, these cases were quite few given the fact that the speakers were given the time they needed to think and compile their ideas before recording. In cases like these, utterances were carefully chosen in a way that such problematic occurrences were avoided for the influence that they might have on the PVI values.

All the recorded data of the reading text and the participants’ spontaneous speech were opened in wideband spectrogram and analyzed using PRAAT software version (5.3.17) by Boersma and Weenink (2012). The vocalic and consonantal interval durations were then calculated using the Pairwise Variability Index. For the sake of obtaining reliable results, double rounds of the measurement were done. The high Pearson product-moment correlation computed ($r=0.93$) implied reliability of the measurements.

3.6.1 Analysis of vowels

Prior to the invent of the spectrogram, very little information has been revealed about the qualities of speech sounds, which was subject to the prejudiced observations of human beings. However, after the availability of the spectrogram, it has become much easier to recognize the properties of speech sounds, particularly the vowels, through the distribution of their formants, (Clark *et al.*, 2007, p.264). According to Kent and Read (2002, p. 110), formant pattern, spectrum, duration and fundamental frequency are the parameters that are relevant to the analysis and description of vowel sounds. However, most of the strategies followed in the acoustic analysis of speech sound, rely on the formant pattern parameter as the primary indication in observing vowels.

Vowels are considered to carry major significance in any speech output due to the fact that they perform dual functions of segmental and suprasegmental tasks. In other words, they can carry information as phonetic segments as well as conveying prosodic information, (Kent and Read, 2002, p 116). However, in the current study, the focus is on its phonetic or segmental information.

The distinct waveforms of vowels that appear on spectrogram make the vowel durations much easier to measure than consonants. That is to say, the vowels acoustic properties are clearer due to the dark striations that mark the inherently greater intensity of the formants. Kent and Read mentioned that the simplicity of the acoustic description and analysis of vowels lies in their ‘steady-state acoustic pattern’ and stable order of the representation of their articulation on spectrogram, (2002, p. 105). And for the phonetic purposes, only the

first three formants are considered as significant, whereas the higher formants have no linguistic indications except that they merely indicate the quality of the speaker's voice, (Ladefoged and Johnson, 2011, p. 195). The first and second formants are of primary significance in terms of the acoustic properties of vowels and thus, for marking the border lines of the start and end points of the vowel sounds. Usually, the first formant, which is considered as the strongest formant, changes with the height of the tongue and the second formant changes with the progress of the tongue.

The interval durations of vowels, according to Clark *et al.* (2007, p. 268) can consistently be taken 'through the stable spectral structure' of the unchanging formants that parallel with the time axis, that is to say, from the beginning to the end of the dark bars that mark the boundaries of the range for the formant frequencies, (Ladefoged, 2001, p. 38). Thus, for the analysis of this study, the interval durations of vowels were consistently taken from the beginning to the end of the dark bars that mark the boundaries or are indicative of the start and end points of the vowel sound. Figure 3.3 shows an instance of spectrographic analysis of the vowel /aa/ produced by a speaker of the present study.

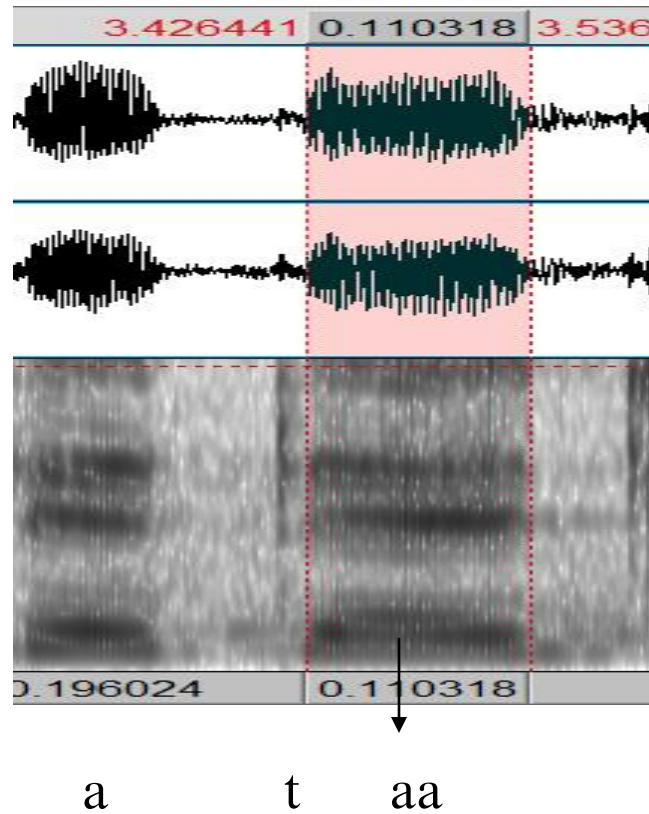


Figure 3.3: Spectrogram of the production of the sound /aa/ by Jordanian speaker 5

All the vowel durations were then measured using the normalized version of the Pairwise Variability Index (PVI):

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

Where ‘ m ’ refers to the ‘number of items in an utterance’ whereas ‘ d ’ refers to the ‘duration of the k th item’, (Grabe and Low, 2002, p. 3). The index calculates the duration differences of successive pairs of vowels, dividing the absolute value of the difference by the mean duration of the pair.

3.6.2 Analysis of Consonants

3.6.2.1 Stops

Whereas the vowels have somewhat similar characteristics, consonants display significant differences in their acoustic properties, (Kend and Read, 2002, p. 139). And since the articulatory organs are continuously moving in speech resulting sometimes in ‘simultaneous’ or ‘overlapping’ articulations, the acoustic properties of speech output display changeable landmarks, (Clark *et al.*, 2007, p. 278). Whereas vowels are usually characterized by simple and uncomplicated landmarks, consonants are considered to have more complex signals on spectrogram. The consonants are, as Ladefoged (2001, p. 47) pointed out, are means of initiating or ending a vowel sound which is typically observed in the voiced stops /b/, /d/, /g/ and /q/. The lips (in the case of /b/), the tongue (in the case of /d/ and /g/) and the uvula (in the case of /q/) undergo a rapid movement before or after a vowel sound. As such, the formant frequencies display an upward movement from these voiced consonants to the higher frequencies of the vowel formants, and then appears a downward movement of formants to the lower formant frequencies of the ending consonant. It is worth mentioning here that the /g/ sound does not exist in SA, however, it occurs in YAD as a replacement of the sound /dʒ/ and in JAD as a replacement of the sound /q/, thus it was measured only in the utterances of Yemeni and Jordanian speakers in recording their spontaneous speech. In the case of the voiceless stops, the /p/ sound does not occur either in SA, JAD or in YAD, thus, only (t and k) occur in the language. The difference between the voiced and the voiceless stops appears when the voiceless stops begin the vowel, where we find they are pronounced accompanied with a slight burst of air and thus producing a different kind of sound. This burst is caused by the occurrence of a period of aspiration between the release of the constriction of the stop sound and the voice onset time of the

vowel. Accordingly, the upward movement of frequencies is less observable in the beginning of the vowel while the downward frequencies that appear in the ending of the vowel are slightly comparable to those of the voiced stops. Generally, stops are characterized by a lower frequency of the first formant. Clark *et al.* (2007, p 281) refer to the formant movements as ‘formant transitions’ and they are quite significant in the identification of consonants in the sense that they represent the rapid movement of the articulatory organs from vowels to consonants or vice versa.

Figure 3.4 displays the segmentation of the voiced and voiceless stop sounds /k/ and /d/:

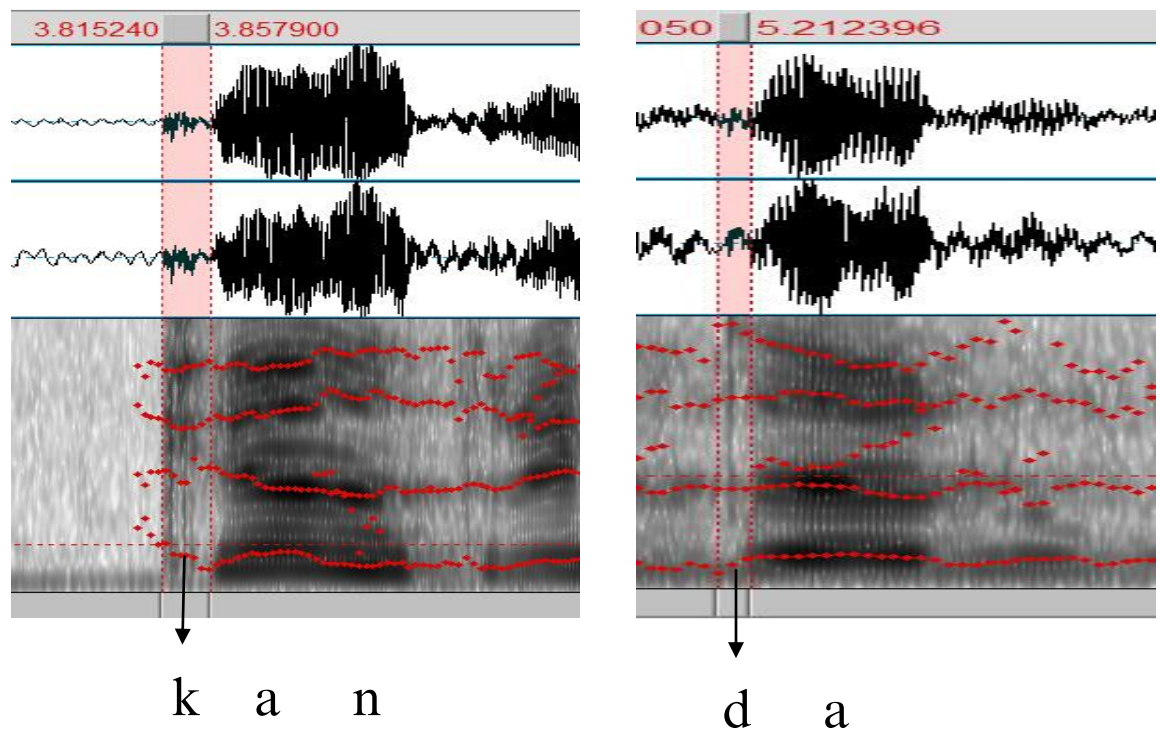


Figure 3.4: Spectrogram of the production of the sounds /k/ and /d/ by Jordanian speaker 1

Being a language that consists also of the glottal stop ‘ʔ’, this sound had to be identified and measured for this study. They were regarded as the most difficult kind of stops to be

measured due to the fact that the glottal stop does not display formant transitions comparable to those seen in the other stop consonants. However, it appears on spectrogram as an interruption of voicing represented by an interval of lowering in the energy that is associated with a sudden vibration of the vocal folds at the beginning of the vowel, (Kent and Read, 2002, p. 185).

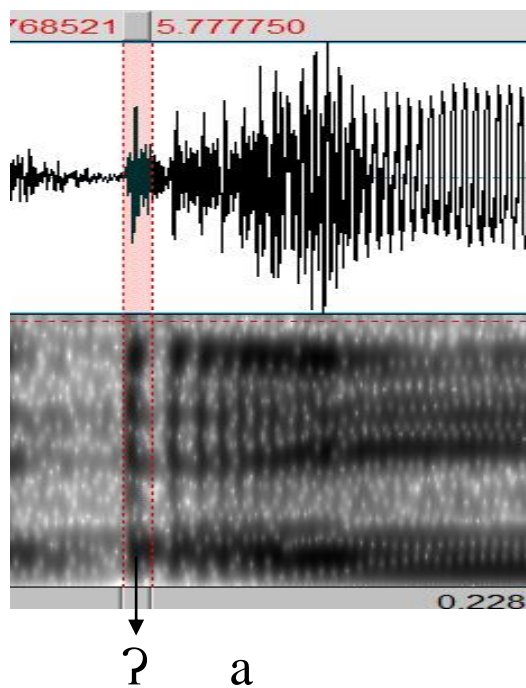


Figure 3.5: Spectrogram of the production of the sound ‘ʔ’ by Jordanian speaker 2

3.6.2.2 Fricatives and Affricates

In the case of fricatives like /f/, /s/, /z/, /θ/, /ð/, /ʃ/, /x/,/ç/, /ħ/ and /h/, the formant transitions are less apparent than those found in stops, however, vertical striations appear on the spectrogram and these striations are slightly darker in fricatives like /s/, /z/, and /ʃ/ due to their louder production and greater intensity. With regard to fricative sounds like /h/,

Ladefoged (2001, p. 57) pointed out that the resonance of the entire vocal tract appears more clearly due to the fact that the sound is produced in the deep vocal tract at the pharynx and glottis, and thus displaying similar characteristics or similar shapes of the vocal tract as those of vowels,. However, the voiced fricatives /z/, /ð/, /x/ and /ʁ/ appear to be produced with vibration of the vocal folds that are observed on spectrogram as discontinuous or interrupted vertical lines.

The voiced affricate /dʒ/ is also displayed on spectrogram as vertical dark striations that involve vibration of the vocal folds that can be seen through the interrupted vertical lines. As for the voiceless affricate /tʃ/, it does not exist in the consonantal system of Arabic. Figure 3.6 shows the segmentation of the fricative /x/ and affricate /dʒ/:

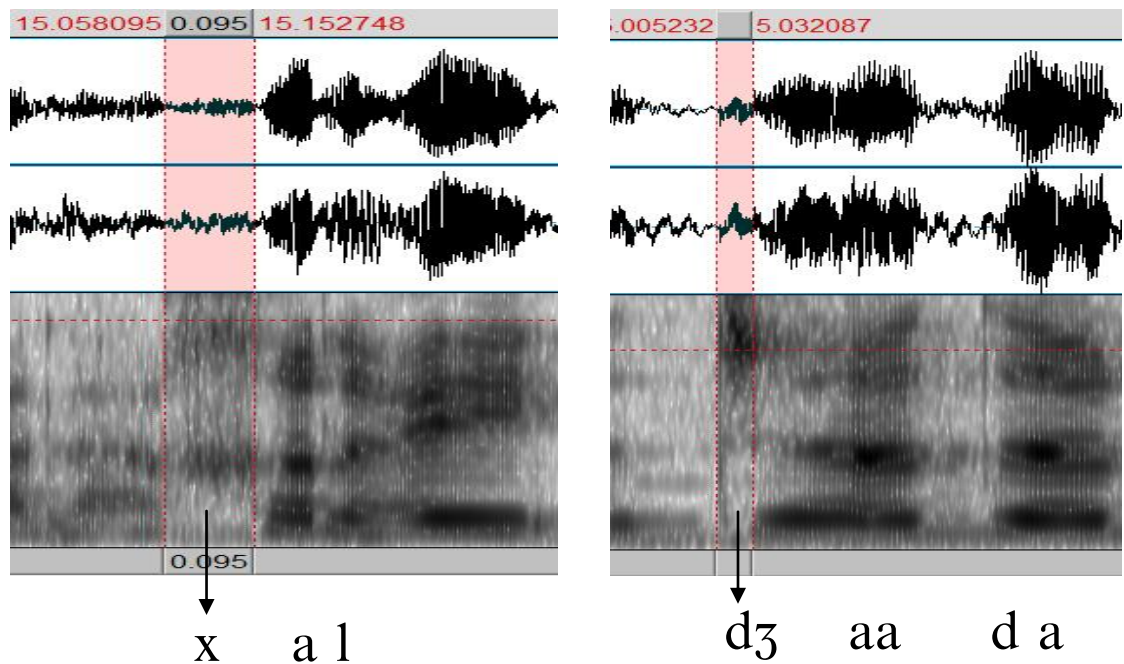


Figure 3.6: Spectrogram of the production of the sounds /x/ and /dʒ/ by Jordanian speaker 4

3.6.2.3 Glide Consonants

The glide consonants /w/ and /j/ are characterized by a gradual movement of the articulatory organs, that is why they are also referred to as semivowels for standing mid way between ‘stops and vowel-vowel transitions’, (Kent and Read, 2002, p.177). Generally, the glide consonants have low-frequency energy. These sounds are considered to be quite problematic in their identification. Their identification in the study was consistently based on taking the middle point between both the glide consonant and the adjacent vowel. Figure 3.7 shows how the glide consonant was captured in this study:

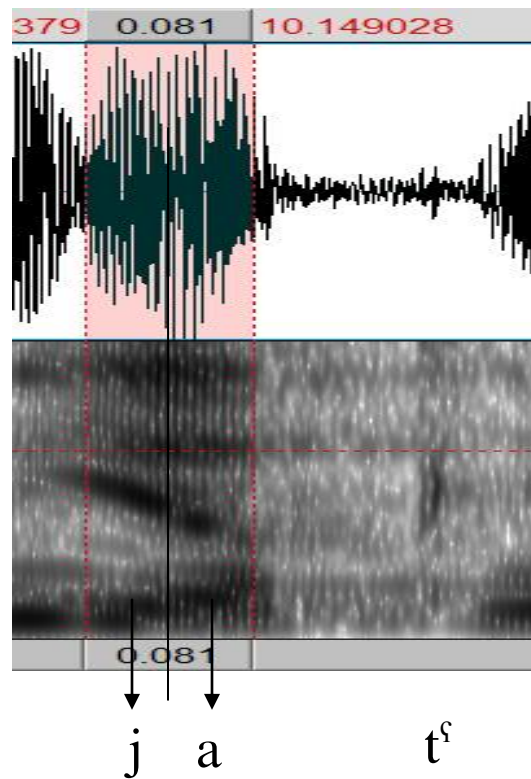


Figure 3.7: Spectrogram of the production of the sound /j/ by Yemeni speaker 2

3.6.2.4 Liquid Consonant /l/ and Trill /r/

Like the glide consonants the liquid consonant /l/ has low-frequency energy. As for the Arabic trill /r/ it displays different signals than the English liquid /r/ sound, where it appears in spectrogram in the form of interrupted vertical lines. Thus, it was measured by marking the start and end points of these vertical lines. Figure 3.8 shows the spectrographic analysis of these two sounds.

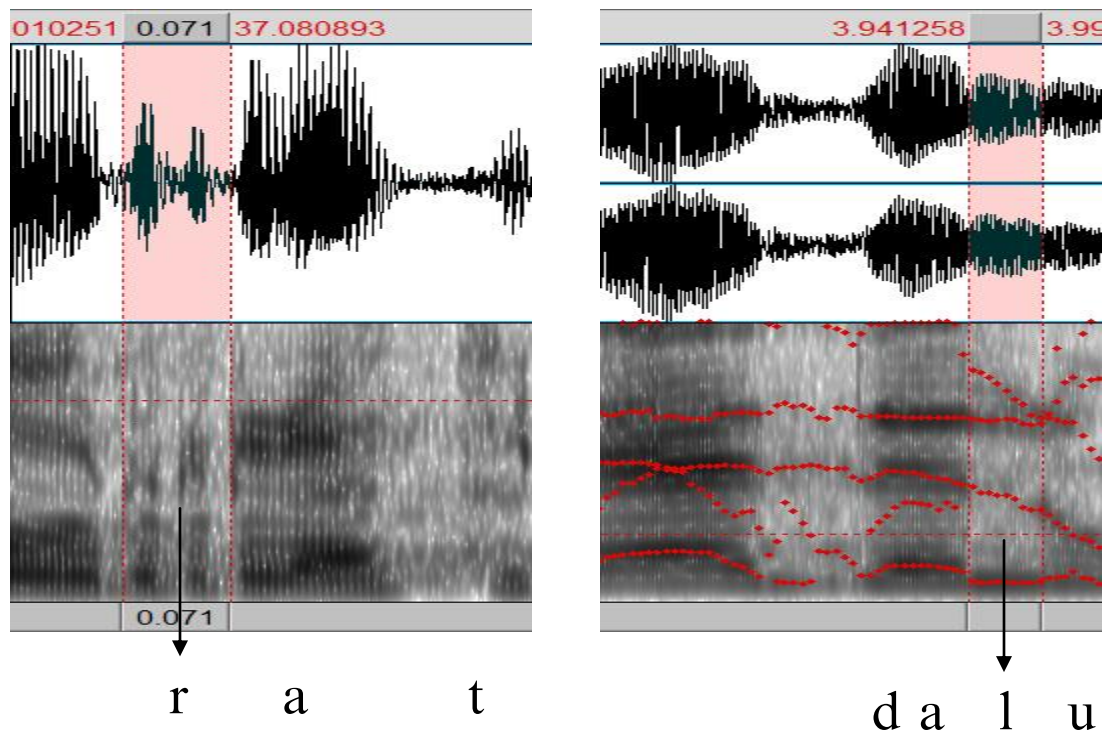


Figure 3.8: Spectrogram of the production of the sounds /r/ and /l/ by / by Yemeni speaker 2

3.6.2.5 Nasal Consonants

The Arabic nasals are restricted to /m/ and /n/ sounds and the /ŋ/ sound is not considered as part of the phonological system of the language. However, we do find some instances

where the /ŋ/ sound is produced as a result of the coarticulation of /n/ if followed by one of the following sounds: /sʰ/, /s/, /θ/, /ð/, /k/, /dʒ/, /ʃ/, /q/, /d/, /t/, /z/, /f/, /t/, /ðʰ/ and /dʰ/. But this is strictly limited to the recitation of the Holy Qura'an and not other texts. As such, only the sounds /m/ and /n/ are analyzed in this study as the nasal consonants in Arabic. Generally, the nasal consonants are produced as a result of total blockage of the vocal tract and the flow of air from the nose. They resemble the vowels in that they can be easily identified through their formant frequencies, however, the difference lies in the louder formants of the vowels. This makes the nasal consonants display low energy of the first formant, (Ladefoged, 2001, p. 53). Thus, the nasals /m/ and /n/ were identified in this study through their clear formant transitions as shown in figure 3.9:

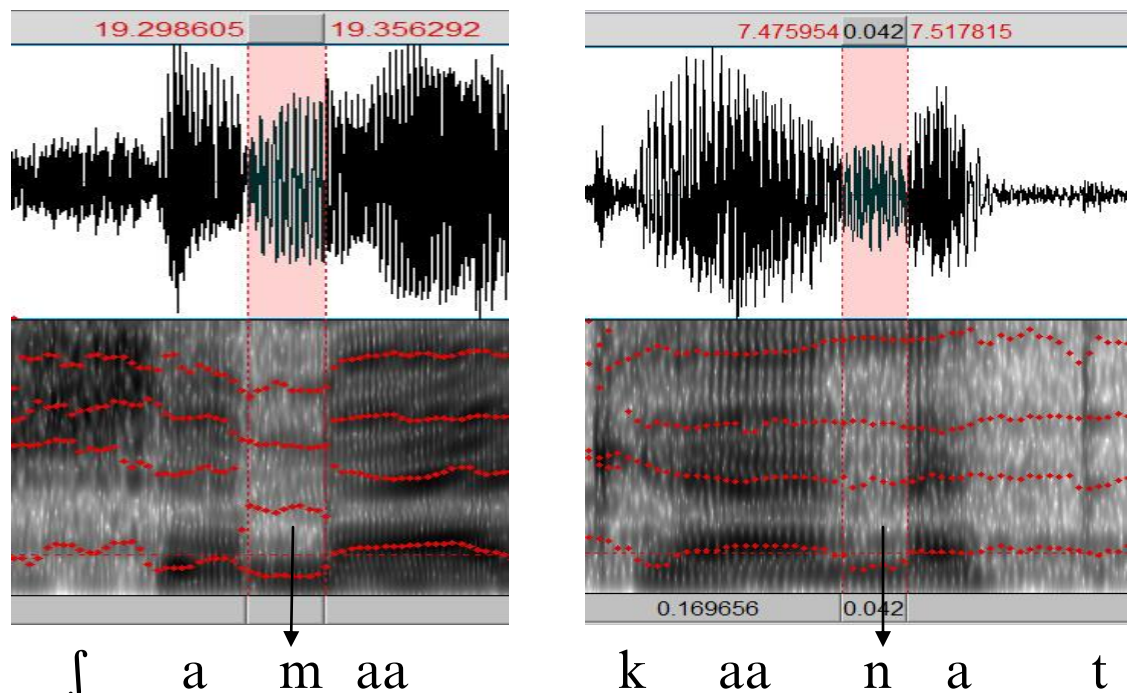


Figure 3.9: Spectrogram of the production of the sounds /m/ and /n/ by Yemeni speaker 1

3.6.2.6 Pharyngealized Consonants (Emphatics)

The emphatics /tˤ/, /ðˤ/, /ʕ/, /dˤ/· /sˤ/ and /lˤ/ are the most distinguished consonant sounds that characterize the Arabic language. Grammarians are particular in using the term ‘emphatics’ or ‘pharyngealized sounds’ to distinguish them from pharyngeal sounds. In the case of pharyngeal sounds, the pharynx is the main place of articulation, whereas in the case of the pharyngealized sounds, the pharynx is the secondary place of articulation. The Arabic terms that are used to describe these sounds are *itbaq* ‘covering’ or *istilaa* ‘elevation’, (Al-Masri and Jongman, 2003). Both terms describe the rising of the tongue and then its backward movement and diffusion all over the pharynx. Thus, the constriction occurs in the upper pharynx, whereas the constriction in the pharyngeal sounds (not the pharyngealized) occurs below the epiglottis. Shar and Ingram (2010, p. 5) mentioned that pharyngealized sounds influence the adjacent vowel by causing a raising of the F1 and the lowering the F2.

Generally, in cases where there were unclear acoustic cues of the sounds, the researcher was able by listening repeatedly to the sounds, to mark the beginning and ending point of these sounds. A second round of the analysis was done for all the sounds produced to confirm the consistency in the segmentation process and the reliability of the measurements.

Figure 3.10 shows the identification of the sound /sˤ/ which displayed dark striations similar to the fricative sound /s/.

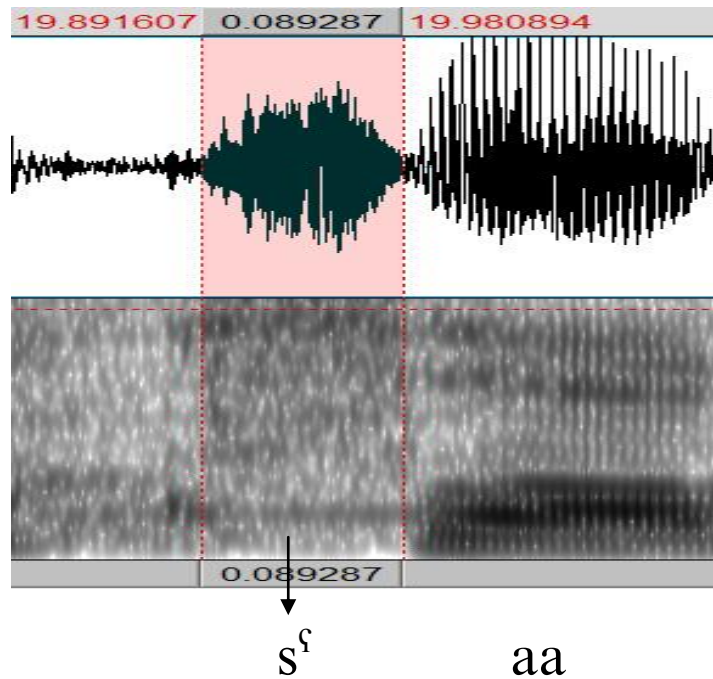


Figure 3.10: Spectrogram of the production of the sound /sʕ/ by Jordanian speaker 1

All the consonantal durations were then measured using the raw Pairwise Variability Index:

$$rPVI = \left[\sum_{k=1}^{m-1} |d_k - d_{k+1}| / (m - 1) \right]$$

This formula gives the sum of the differences between successive pairs of consonantal measurements divided by the total number of those measurements in an utterance minus one.

3.7 Conclusion

This chapter introduced the research design and the analytical framework of the study. It provided information on how the PVI was used to analyze the recorded data in terms of the durational intervals of the vowels and consonants in order to help in describing the language in question. Detailed descriptions of the materials, participants, data collection and data analysis were provided.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

As mentioned earlier in chapter 1, the study compares between YAD and JAD by analyzing both the vocalic and consonantal durational intervals. The following sections present the results of the measurements of both the consonantal and vocalic durations of SA read text and spontaneous speech utterances of the Yemeni and Jordanian speakers. All the values are presented and analyzed in order to account for the rhythm of YAD compared to that of JAD. Dauer's (1983) criteria are also discussed in the study to find how the phonological features of the dialects influenced their consonantal and vocalic durations and thus influenced the rhythm output.

4.2 Reading Text (SA)

4.2.1 Consonantal Durations of Yemeni and Jordanian Speakers

The measurements obtained from each Yemeni and Jordanian speaker's utterances of *The North Wind and the Sun* text, were investigated for significant differences. A total of eight utterances were extracted and they were identified by expressing a single intonational phrase. Out of these utterances, 126 consonantal durations were measured for each Yemeni and Jordanian speaker resulting in a total of 1,260 consonantal durations measured for all the ten speakers of both dialects. Having measured the consonantal duration using the raw

version of the PVI, this section presents the *r*PVI values obtained from the Yemeni and Jordanian speakers' utterances of the read text. The average *r*PVI values of the consonantal durations are reported in Table 4.1 and 4.2:

Table 4.1: Computed *r*PVI values of the consonantal durations of the SA read text of Yemeni speakers

Speaker	Yemeni <i>r</i>PVI values
1	38.6
2	36.16
3	47.91
4	36.56
5	37.83
AVERAGE	39.40

Table 4.2: Computed *r*PVI values of the consonantal durations of the SA read text of Jordanian speakers

Speaker	Jordanian <i>r</i>PVI values
1	36.98
2	40.60
3	52.90
4	38.54
5	35.94
AVERAGE	40.99

4.2.2 Vocalic Durations of Yemeni and Jordanian Speakers

This section presents the results of the measurements of the vocalic durations of the Yemeni and Jordanian speakers' utterances of the SA read text. Out of the eight utterances extracted for analysis, 102 vowel durations were measured for each Yemeni and Jordanian speaker, resulting in a total of 1,020 vocalic durations measured for the ten speakers from both Arabic dialects. As mentioned in section 3.2, the *n*PVI is used for measuring the vocalic durations due to the vowels' sensitiveness to speech rate. The *n*PVI values of the vocalic durations and their average proportions obtained from the Yemeni and Jordanian speakers' recordings are given in Tables 4.3 and 4.4:

Table 4.3: Computed *n*PVI values of the vocalic durations of the SA read text of Yemeni speakers

Speaker	Yemeni <i>n</i>PVI values
1	54.78
2	46.11
3	55.11
4	48.87
5	48.84
AVERAGE	50.74

Table 4.4: Computed *n*PVI values of the vocalic durations of the SA read text of Jordanian speakers

Speaker	Jordanian <i>n</i>PVI values
1	45.17
2	53.57
3	47.58
4	56.16
5	54.08
AVERAGE	51.29

These results can be illustrated in figures 4.1 and 4.2:

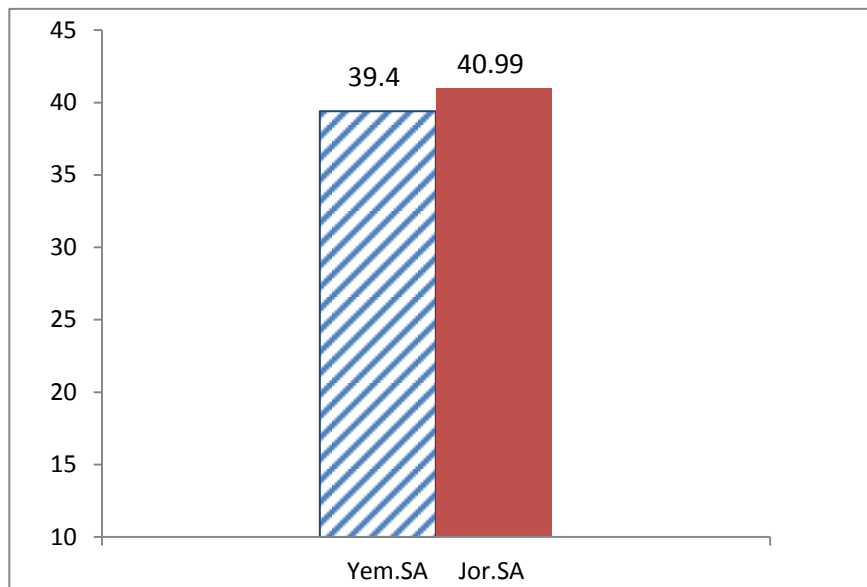


Figure 4.1: Computed *rPVI* values of the consonantal durations of Yemeni and Jordanian Speakers of the read text

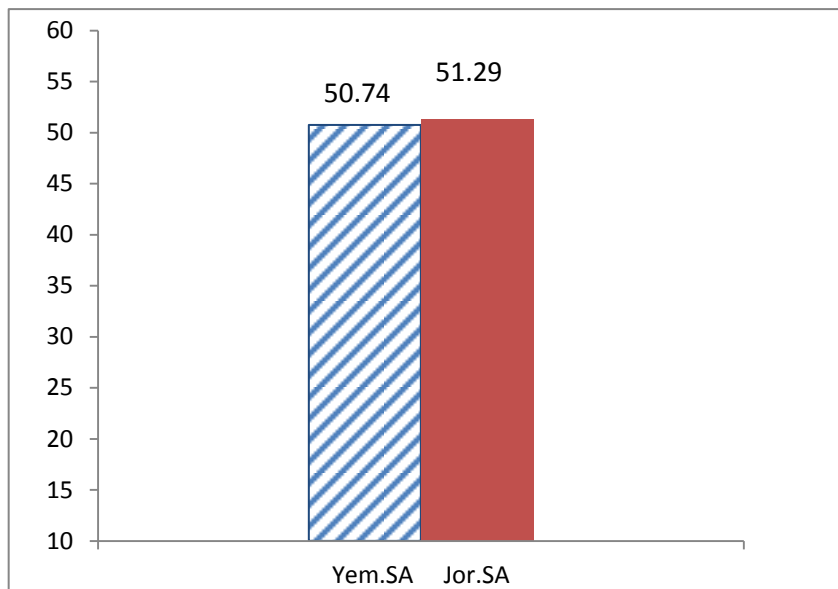


Figure 4.2: Computed *nPVI* values of the vocalic durations of Yemeni and Jordanian Speakers of the read text

Figure 4.1 presents the r PVI values of the average proportions of the consonantal durational intervals obtained from the Yemeni and Jordanian speakers' recordings of the SA read text. Given the fact that both Yemeni and Jordanian speakers produced eight utterances of a fixed Standard written text, a paired t -test was used to compare between the PVI values of each pair of Yemeni and Jordanian speakers' utterances. The results of the t -test showed no significant difference between the two sets of the SA utterances produced by the speakers of YAD and JAD ($t(7) = 1.80, p = 0.11$).

Figure 4.2 presents the average n PVI values obtained from the durations of successive pairs of vowels occurring in SA text utterances. Like the PVI values of the consonantal durations, the PVI values of the vocalic durations of each pair of Yemeni and Jordanian speakers' utterances were also compared using a paired t -test showing no significant differences between the two sets of utterances ($t(7) = 0.39, p = 0.73$).

4.3 Spontaneous Speech

4.3.1 Consonantal Durations of YAD and JAD Speakers

The utterances of the spontaneous speech were selected from the participants' conversations regarding a topic that they were interested in (see section 3.5.2). Similar to the method used for the identification of utterances for the read text, identification of utterances of spontaneous speech was done through the pause locus. Given the naturally long sentences in Arabic language, the number of the syllables in the selected utterances ranged between ten to twelve syllables. 645 consonantal durations were measured for Yemeni speakers and 660 consonantal durations were measured for Jordanian speakers.

Thus, a total of 1,305 consonantal durations were measured for the ten speakers from both dialects. The average *r*PVI values of the consonantal durations are listed in Tables 4.5 and 4.6:

Table 4.5: Computed *r*PVI values of the consonantal durations of YAD spontaneous speech

Speaker	YAD <i>r</i>PVI values
1	38.23
2	37.14
3	39.29
4	34.21
5	44.29
AVERAGE	38.63

Table 4.6: Computed *rPVI* values of the consonantal durations of JAD spontaneous speech

Speaker	JAD <i>rPVI</i> values
1	47.86
2	35.71
3	41.15
4	38.57
5	42.67
AVERAGE	41.19

4.3.2 Vocalic Durations of YAD and JAD Speakers

Out of the eight utterances extracted from the YAD and JAD spontaneous speech, a total of 890 vocalic intervals were measured for the Yemeni and Jordanian speakers. 441 vocalic durations were measured for the Yemeni speakers and 449 vocalic durations were measured for the Jordanian speakers. The *nPVI* values of the vocalic durations are reported in Tables 4.7 and 4.8 and figures 4.3 and 4.4:

Table 4.7: Computed *n*PVI values of the vocalic durations of YAD spontaneous speech

Speaker	YAD <i>n</i>PVI values
1	51.00
2	40.13
3	49.11
4	64.33
5	67.73
AVERAGE	54.46

Table 4.8: Computed *n*PVI values of the vocalic durations of JAD spontaneous speech

Speaker	JAD <i>n</i>PVI values
1	62.49
2	49.32
3	55.87
4	69.86
5	54.58
AVERAGE	58.42

These results can be illustrated in figures 4.3 and 4.4:

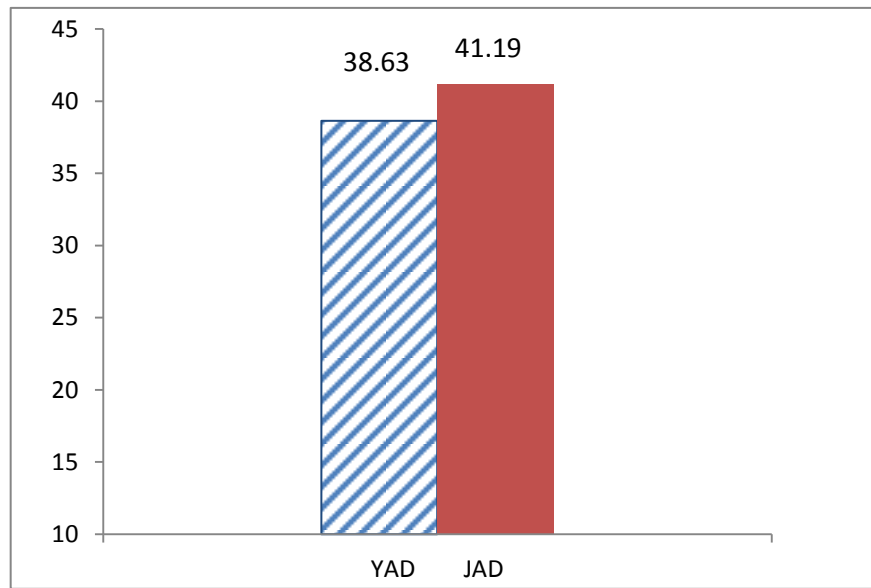


Figure 4.3: Computed *rPVI* values of the consonantal durations of YAD and JAD Speakers in spontaneous speech

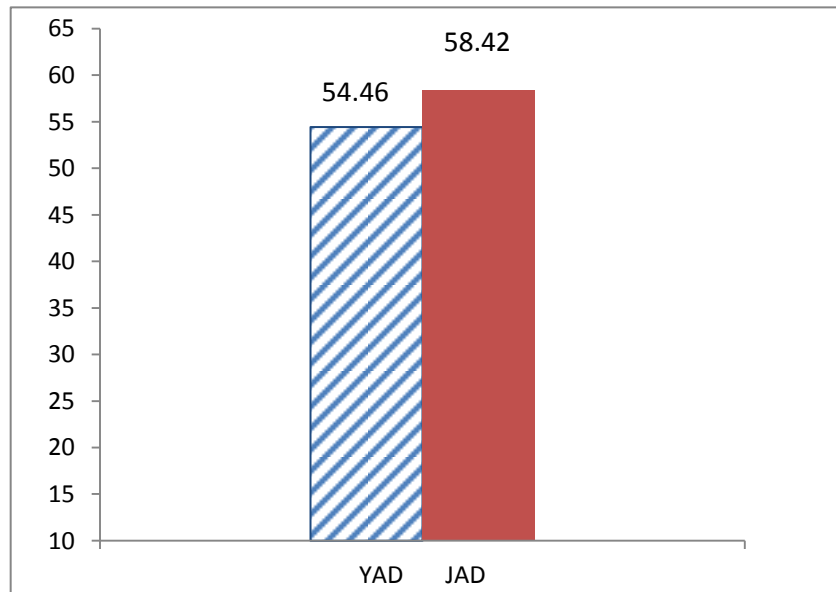


Figure 4.4: Computed *nPVI* values of the vocalic durations of YAD and JAD Speakers in spontaneous speech

Figure 4.3 presents the average r PVI values of the consonantal durations in YAD and JAD spontaneous speech. An independent t -test (equal variance) was used to compare between the PVI values of consonantal durations of the Yemeni and Jordanian speakers. The results of the t -test showed no significant differences between the utterances of YAD and JAD ($t(8) = 0.97, p = 0.35$). The effect size value ($r = 0.326$) suggested a moderate practical significance.

Figure 4.4 presents the average n PVI values of the vocalic durations of spontaneous speech in both YAD and JAD. The results were also compared using an independent t -test showing no significant differences between the durations of the successive vowels in both dialects ($t(8) = 0.63, p = 0.541$). The effect size value ($r = 0.22$) suggested a low practical significance. The results, thus, indicate that the two Yemeni and Jordanian dialects do not pattern differently in terms of the consonantal durational intervals.

4.4 Discussion

As can be seen from the above results, no significant differences were found between the YAD and JAD consonantal and vocalic durational values whether in the SA read text or even the spontaneous speech with a probability values ($p > 0.05$). Both dialects displayed comparable values that indicate similar rhythmic structures in these Arabic dialects. Hamdi *et al.* (2004) suggested that the rhythmic variations between dialects or languages are strictly proportional to the distance between the regions where these dialects or languages are spoken. This can be inferred from the authors' comparable results of the Arabic dialects that appeared in the east like Jordanian and Syrian; and the highly significant differences that appeared when compared with the western Moroccan dialect, showing a probability

value of ($p < 0.001$). However, in this study, the geographical divergence seemed to have little influence over the rhythmic productions of the Yemeni and Jordanian speakers since no significant differences appeared between their computed values despite the distance between the countries where these Arabic dialects are spoken. The homogeneity of the participants of the study in terms of their gender and educational background where all of them were postgraduate female students could be a possible factor contributing to the similarity of their rhythm output. In investigating the use of the Arabic dialects in the Arabic communities, Harry suggested that the speaking skills of the speakers are directly linked to their level of education. The highly educated people seem to display similar speaking skills which tend to be quite formal and acrolectal since they use an elevated or semi-literary colloquial rather than the plain colloquial which is used by the illiterates or even the less educated speakers (1996, p. 76). Thus, it can be said that since all the participants in the current study are postgraduate students, they spoke a similar acrolectal form that resulted in similar rhythmic output of speakers of both dialects. It should also be noted that this study has been conducted on participants in a non-native context. Although the researcher attempted to cautiously choose speakers who had spent less than a year in Malaysia to avoid influences over their rhythm output, there is a possibility for Arabic speakers of different dialects to reach at comparable speech and rhythm output when living in a foreign and non-Arabic country where their dialects might intermingle. This could also be a possible factor that affected the findings of the study.

The current study reveals that YAD and JAD do not seem to display many variations in the consonantal durations with relatively low PVI values obtained from both the read text and spontaneous speech utterances as can be seen in Tables 4.1, 4.2, 4.5 and 4.6. This could

possibly be attributed to the less varied and less complex syllable structures that do not allow for many consonant clusters which Dauer (1983) considered as a contributing factor to the rhythm patterning of a language. According to Dauer, the variety of syllable structures and syllable complexities enhance the stress timing features of a language like English that displays different syllable types and sequences of consonants. In this study, however, the speakers' utterances in both YAD and JAD do not seem to display such variations in the syllable structure where the most frequently occurring syllable type is CV:

YEMENI	JORDANIAN
CV (60%)	CV (58%)
CVC (37%)	CVC (34%)
CVCC (3%)	CCV (8%)

This rareness of diverse and complex syllable structures in both dialects could be a possible reason for the low PVI values of the consonantal durations which supports Dauer's claim that the less varied syllable types in a language yields less stress-timing features in that particular language. Moreover, the similarity of the syllable structures of both YAD and JAD could explain the similarity of their rhythm output or the insignificant differences between the *r*PVI values obtained from the consonantal durations measurements of both dialects.

It is worth mentioning that the above structures show that YAD and JAD display some differences in the syllable structure. Although investigating the syllable structures of YAD and JAD was beyond the scope of the study, it has been observed that the Jordanian

speakers appeared to simplify the consonant clusters in the word final position by a process of vowel insertion. Such words as ‘subḥ’ (morning), ‘ʔakl’ (food), ‘sabt’ (Saturday), ‘baḥr’ (sea) and ‘ʃaribt’ are produced in JAD with a simplification of the final consonant cluster by means of a vowel insertion resulting in the realization of these words as ‘subuḥ’, ‘ʔakil’, ‘sabıt’, ‘baḥır’ and ‘ʃrıbit’ respectively. In short, whereas JAD appeared to allow for a cluster consisting of two consonants word initially only, YAD appeared to allow only for cluster of two consonants word finally only.

As for the vocalic durations, both YAD and YAD displayed high PVI values of the intervals of vocalic durations both in text reading and spontaneous speech as can be seen in Tables 4.3, 4.4, 4.7 and 4.8. These high values indicate stress timing features in terms of vowels in both dialects. The following figure shows an instance of vowel occurrences in long and short contrasts which might have possibly been captured by the PVI resulting in high values that reflect variational vocalic intervals. Note that ‘v’ denotes a short vowel while ‘V’ denotes a long vowel.

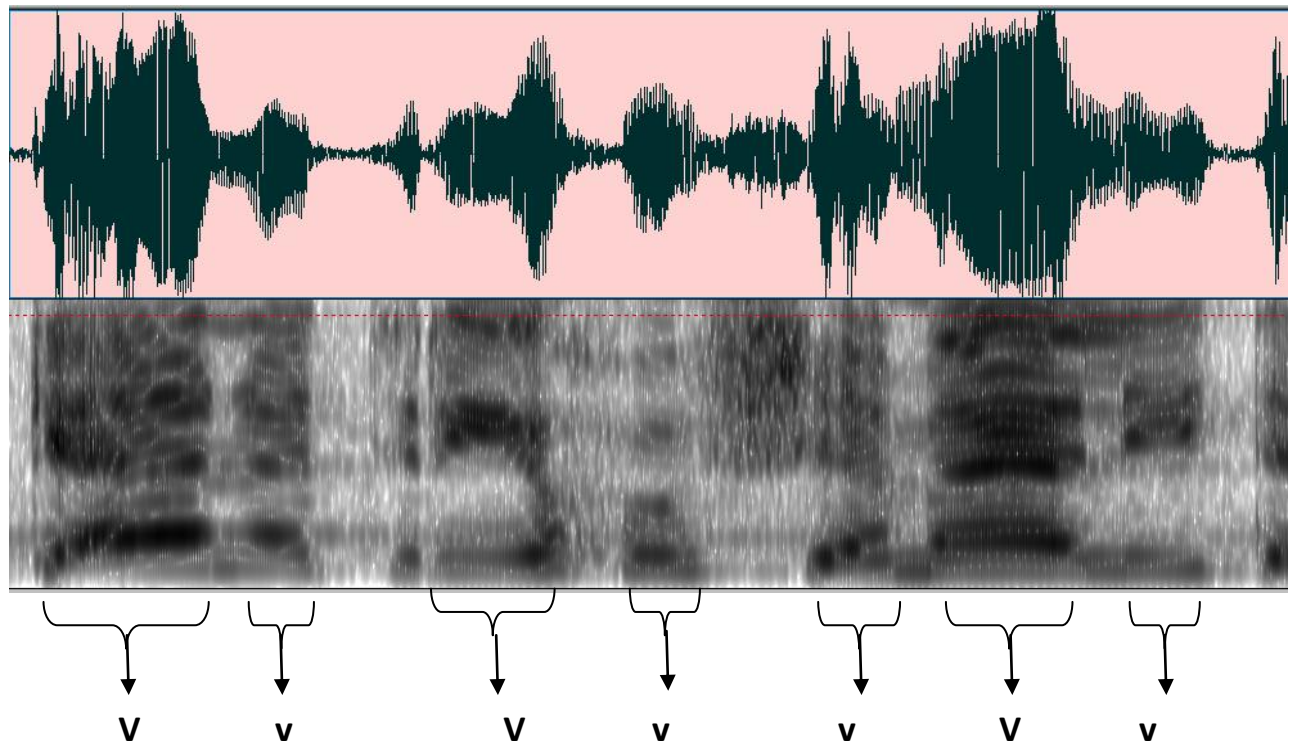


Figure 4.5: Spectrogram of the vowel contrasts in Arabic

At this stage, one can assume that the phonological features of a language such as the syllable structure or vowel duration can possibly have an influence over the rhythm output, but we cannot be certain of the comprehensiveness of these criteria in reflecting speech rhythm since the indexes could only provide the measurements of the durational variability and not the origin of these variations.

Given the high *n*PVI values of the vocalic durations and the low *r*PVI values of the consonantal durations, one can thus assume that different results and classifications can be obtained by two rhythmic parameters. Consider the following figure that shows the distribution of YAD and JAD over the *n*PVI and *r*PVI scale:

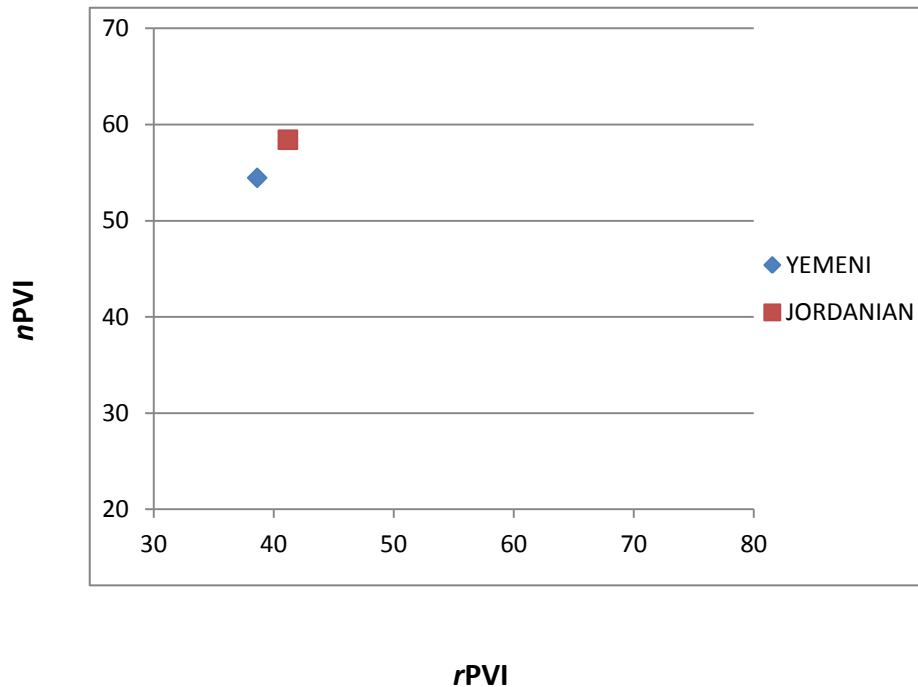


Figure 4.6: Distribution of YAD and JAD and other languages over the *nPVI* and *rPVI* scales

The above figure shows the distribution of YAD and JAD based on the values of the consonantal durations (*rPVI*) in the x-axis and the vocalic durations (*nPVI*) on the y-axis. In comparing the distribution of languages in figure 4.6 to Grabe and Low's distribution of languages that have been prototypically assigned to a particular class of rhythm (see figure 2.2), we find that YAD and JAD do not appear to fit in the rhythm continuum observed in Grabe and Low's distribution. Their distribution demonstrates the nature of rhythm in terms of a continuum where some languages appeared as typical stress-timed languages, some as typical syllable-timed languages and others appearing midway between these two classes in what is referred to as intermediate languages. These results seem to support Dauer's (1983) postulation of rhythm as a continuum where no evidence is found for dichotomous or binary clusters of languages.

In this study, however, both YAD and JAD did not seem to fit in Grabe and Low's continuum. Although YAD and JAD displayed mixed results with high variability of the vocalic durations and low variability in the consonantal durations, Figure 4.6 shows that both dialects do not seem to occupy a position near the position of the prototypical intermediate languages Catalan and Polish distributed in Grabe and Low's figure, rather, they occupied a distinct position that denotes their mixed features based on two parameters which put them in two different classes of rhythm. In terms of the vocalic durations, they appeared close to the prototypical stress-timed English, while in terms of the consonantal duration, they occupied a position much lower than the prototypical syllable-timed Spanish.

Like Estonian in Nolan and Asu's (2009) study which appeared to be rhythmically similar and different from English based on measuring both the syllable and foot, in this study also, the two parameters of the consonantal and vocalic durations lead to different classifications of Arabic. Prior to Nolan and Asu (2009), Nespors (1990) seemed to support the notion of mixed features of rhythm. Catalan and Polish, according to Nespors (1990), are mixed languages where Polish is considered as syllable-timed with fewer complexities in the syllable structure but also involving occurrences of vowel reduction while Catalan is considered as stress-timed due to the higher complexities of the syllable structure but without involving occurrences of vowel reduction. Similarly, YAD and JAD appeared to have some mixed features in that they displayed vowel contrasts that resulted in varied vocalic durations while in terms of the syllable structure, both dialects appeared to have few complexities in their syllable types. However, despite having CV syllable structure as the most prevailing syllable type in the speakers' utterances, other syllable types were also

used such as the CVC (37%) and CVCC (3%) in YAD, and CVC (34%) and CCV (8%) in JAD.

However, although the notion of classes of rhythm seems to be enhanced in this study rather than dichotomous or continuous rhythm, YAD and JAD need to be compared to a wider range of languages to provide a clearer image of their distribution among a number of other languages within the same study to further support or contradict this assumption.

4.5 Conclusion

This chapter reported the results of the measurements of the consonantal and vocalic durational intervals of utterances produced by both Yemeni and Jordanian speakers. It provided a full presentation and discussion of the results obtained from the analysis of the study with reference to the materials used for eliciting data, namely, the read text and spontaneous speech. A comparison was made between the values obtained from SA utterances read by Yemeni and Jordanian speakers and those obtained from YAD and JAD spontaneous speech. The following chapter will conclude the findings of the study with reference to the research questions in order to explain how the objectives set for this study were fulfilled.

CHAPTER 5

CONCLUSION

5.1 Introduction

The previous chapter reported the results of the analysis of the consonantal and vocalic durational intervals of the utterances produced by YAD and JAD speakers both in reading and spontaneous speech. Based on the results obtained from the current study, this chapter provides answers to the research questions stated in section (1.7) in order to find out how the objectives of this study were achieved.

The first part of this summary attempts to answer the first research question by describing the acoustic correlates of the rhythm of SA as read by Yemenis and Jordanians. The second part attempts to answer the second research question by describing the acoustic correlates of the rhythm of YAD and JAD in spontaneous speech. The third part attempts to answer the third research question by comparing the rhythm of YAD to that of JAD which has long been classified as an eastern dialect.

5.2 Summary of Results

This study set out to answer three research questions:

- (i) What are the acoustic correlates of the rhythm of SA as read by Yemenis and Jordanians?
- (ii) What are the acoustic correlates of the rhythm of spoken Yemeni Arabic dialect and Jordanian Arabic dialect?
- (iii) How does the rhythm of Yemeni Arabic dialect compare with the rhythm of Jordanian Arabic dialect?

Table 5.1 summarizes the results of the PVI values obtained from the analysis of the study:

Table 5.1: Summary of the PVI values

	YAD	JAD	YAD	JAD
	(Reading Text)	(Reading Text)	(Spontaneous speech)	(Spontaneous Speech)
Consonantal				
Durations (<i>r</i> PVI)	39.4	40.99	38.63	41.19
Vocalic				
Durations (<i>n</i> PVI)	50.74	51.29	54.46	58.32

5.2.1 The Acoustic Correlates of the Rhythm of SA as Read by Yemenis and Jordanians

The first research question regarding the acoustic correlates of the rhythm of SA as read by Yemenis and Jordanians is answered in this section. The average PVI value of the consonantal durations of SA as read by Yemenis was 39.4 whereas the average PVI value of their vocalic durations was 50.74. The average PVI value of the consonantal durations of SA as read by Jordanians was 40.99 whereas the average PVI value of their vocalic durations was 51.29.

The acoustic correlates of the rhythm of SA have been investigated by analyzing the consonantal and vocalic durational intervals of the speakers' utterances of a reading passage of the translated version of '*The North Wind and the Sun*'. The durations of the consonants and vowels were analyzed via the use of the *rPVI* and *nPVI*, respectively. Those two rhythmic parameters, namely the consonantal and vocalic intervals, lead to two different findings that reflect mixed rhythmic features and might be considered as problematic in the classification of YAD and JAD. In terms of consonantal durations, the *rPVI* values appeared to be low as can be seen in Table 5.1. These low values indicate syllable-timed rhythm. Ramus *et al.* (1999) suggested, in support of Dauer (1983), that the acoustic measurements are a reflection of certain phonological properties that may influence the rhythmic structure of a particular language. Some of these properties are the variations and complexities in the structures of the language syllables, frequent occurrence or higher possibilities of vowel reduction and the correlates of stress. One can thus assume that the rare occurrences of varied and complex syllable structures in Arabic could be a

possible explanation of these low values of the consonantal durations since Dauer suggested that the simple and less diverse syllable structure yields in less stress timing features of the language.

With respect to the vocalic durations of the YAD and JAD utterances, the *n*PVI values appeared to be high as can be seen in Table 5.1. These high values indicate the greater variations in the durations of successive pairs of vowels which reflect the stress timing features in the language. These features could be, according to Dauer's description of stress-timed languages, a possible effect of certain phonological features related to vowel duration. The stress-timed languages are expected to involve frequent occurrences of vowel reduction in unstressed syllables which results in variational vocalic durations. The absence of this feature yields in more regular vocalic durations which can be seen in syllable-timed languages.

Vowel duration, which is regarded as an acoustic correlate of stress, is much influenced by the stress in most stress -timed languages. Arabic is a language that displays stress, (El-Hassan, 1994, p. 209) and the stress could have possibly resulted in the lengthening and prominence of the syllables that received the stress. However, unlike English or German where the stress contributes to the meaning and grammatical functions of the words in the language, in Arabic, stress has no effect on the meaning and grammatical functions of words even if changing the stress position might lead to a distorted pronunciation of the word, (Chentir, Guerti, and Hirst, 2009). In studying the vowel durations of stressed

syllables in Arabic, Kenneth *et al.* (1999) found that stressed syllables display remarkable vowel lengthening which is associated by an increase in the first formant of that vowel.

Generally, in summarizing the results seen in Table (5.1), the rhythmic structure of YAD and JAD appeared to have mixed features based on two parameters, where the measurements of the consonantal and vocalic durations lead to different results. However, it should be noted that Grabe and Low (2002), White and Mattys (2007) and Wiget *et al.* (2010) considered that measuring the vocalic durations provide better results than the consonantal durations in terms of discriminating languages and capturing the cross linguistic variations that occur between them. As has been seen earlier in Figure (2.4) in section (2.5), Grabe and Low (2002) found that the number of the *n*PVI differences is double the number of the *r*PVI values which lead them to suggest that the *n*PVI seems to be more effective in discriminating between languages than the *r*PVI. However, these conclusions were drawn from a study that counted on one speaker for each language. That is why they suggested that more speakers should be involved to examine the validation and effectiveness of the *r*PVI role in capturing the patterns of rhythm of a language and thus, the classification of that language. Asu and Nolan (2006) on the other hand, applied the syllable PVI and foot PVI in measuring the vocalic and consonantal durations of English and Estonian. As has been mentioned in section (2.5), they found that measurements of the syllable and foot PVIs lead to different results from measurements at the level of stress since English that has traditionally been regarded as controlling for foot rather than syllable appeared to show more variations than Estonian at the syllable level and limited variations at the foot level. It can be inferred from the above that one parameter

may not be sufficient to describe the rhythmic structure of a language while a combination of different rhythmic parameters can lead to a more reliable analysis of rhythm.

5.2.2 The Acoustic Correlates of the Rhythm of YAD and JAD in Spontaneous Speech

The second research question regarding the acoustic correlates of the rhythm of YAD and JAD in spontaneous speech is answered in this section. The average PVI value of the consonantal durations of YAD spontaneous speech was 38.63 whereas the average PVI value of the vocalic durations was 54.46. As for JAD spontaneous speech, the average PVI value of the consonantal durations was 41.19 whereas the average PVI value of the vocalic durations was 58.42.

The consonantal and vocalic durations of YAD and JAD spontaneous speech have been measured by means of the *r*PVI and *n*PVI indexes to investigate the acoustic correlates of the rhythm of YAD and JAD. Like the reading text, the values of the consonantal and vocalic durations of spontaneous speech appeared to have different indications according to the stress-timed/syllable-timed classification. As seen in Table (5.1), the *r*PVI values of the consonantal durations appeared to be low which could possibly be attributed to the relatively regular syllable structure that has been observed in the speakers' utterances with the quite frequent occurrences of the CV syllable structure in both YAD and JAD. These low PVI values can lead to the description of YAD and JAD as having syllable timing features in terms of the consonantal durations.

The vocalic durations, on the other hand, resulted in high PVI values which indicate stress timing features of the YAD and JAD rhythmic structure. This might be a reflection of certain phonological features related to vowel duration such as prominence or vowel reduction. Although several acoustic studies have shown that Arabic dialects display vowel reduction that occurs increasingly as one moves from the east to the west, El-Hassan (1994) claimed that the concept of vowel reduction is alien to Arabic and that both the short and long vowels are fully articulated whether in stressed or unstressed syllables (1994, p. 212). However, in this study, contrasting vowel lengths have been observed in the speakers' utterances which could possibly contributed to the varied vocalic durations reflected by the high PVI values. Accordingly, it might not be convincing to determine currently whether or not vowel reduction is a process that causes the stress timing features of YAD and JAD, especially since Arvaniti (2012, p. 4) suggested that Dauer's postulation that the rhythm of languages can be determined by means of a binary opposition in terms of the absence or presence of the phonological features mentioned in Dauer (1983), might not always be adequate in describing the rhythm of a language. This denotes a possibility of the involvement of other processes that can have an influence over the rhythm of a language.

Generally, given the low *rPVI* values and high *nPVI* values, or in other words the relatively regular consonantal durations and great variations in the vocalic durations, the rhythmic structure of YAD spontaneous speech can be described, like read SA, as having mixed features represented by vowel contrasts and few complexities of syllable structure. Although some researchers described the *nPVI* as being a better index in discriminating between languages, the very low *rPVI* values obtained in this study that appeared to be lower than the *rPVI* values of the prototypical syllable-timed languages, do carry some

indications and significance that enhance its role in capturing the patterning of rhythm in a language.

Besides conforming with Nespor (1990) and Nolan and Asu's concept of mixed features, the findings of this study also coincides with those of Frota and Vigario (2001), where both European Portuguese and Brazilian Portuguese in their study appeared to have mixed rhythm. The former tended to pattern after stress-timed languages in terms of the consonantal durations but is similar to syllable-timed languages in terms of the vocalic durations. Brazilian Portuguese, on the other hand, patterned after syllable-timed languages in terms of the consonantal durations but resembled the mora-timed languages in terms of the vocalic durations. These results enhance the concept of classes in rhythm typology indicating that intermediate rhythm and mixed or coexisting rhythm are not equivalent.

5.2.3 YAD Rhythm Compared to JAD

This section answers the third research question by comparing YAD rhythm to that of JAD. Like YAD, the consonantal durations of JAD resulted in low PVI values, which like YAD rhythm, indicates syllable timing features of the JAD rhythmic structure. The vocalic durations, on the other hand, resulted in high PVI values which indicate stress timing features of both YAD and JAD rhythm. However, it has to be pointed out that if only the *n*PVI values are considered in this study, the results would have lead to the classification of both YAD and JAD as stress-timed dialects.

It has been mentioned in chapter 1 that YAD and JAD display some differences in certain phonological features such as the syllable structures and some vowels. These variations were expected to have some influences over the rhythm of both dialects. Some instances of syllable structure differences have been observed in the data of this study in the sense that YAD allows only for a final cluster of two consonants whereas Jordanian speakers tended to simplify the final cluster by means of vowel insertion and thus allowing for a cluster of two consonants word initially only. Although the variations between YAD and JAD phonological features were expected to be reflected by significant differences between the PVI values obtained from the utterances of speakers from both dialects, the results of the study contradicts these expectations since both dialects displayed comparable intervals of consonantal and vocalic durations. The results of the *t*-tests showed no significant differences between the Yemeni's and Jordanian's utterances both in SA read text and spontaneous speech.

Both Hamdi *et al.* (2004) and Ghazali *et al.* (2002) investigated the rhythm of the Arabic dialects using the Ramus *et al.* (1999) ΔC and %V. Hamdi *et al.* (2004) found that the values of ΔC increased as they moved from the east to the west while the %V decreased, showing that JAD occupied a position with eastern Arabic dialects with less stress timing features, quite distinctly from Moroccan and Algerian dialects that appeared as western Arabic dialects with highly stress-timed features.. Similar results were found by Ghazali *et al.* (2002) in the sense that the North African Arabic dialects displayed high stress timing features with a much higher ΔC and lower %V compared to eastern Arabic dialects like Jordanian and Syrian. The probable deletion of the reduced vowels in the North African dialects was considered as a highly contributing factor to their low %V values, whereas the

ΔC indicated the changeable number of consonants that was triggered by this deletion of the reduced vowels and thus manifested by different intervals of consonantal durations in these dialects. Similar complexities were not recognized in the eastern Arabic dialects although they were also classified as being stress –timed. The current study revealed that YAD and JAD do not display recognizable variations in the consonantal and vocalic durations with no significant differences between the two dialects in the read text and spontaneous speech. One can thus assume that based on the previous classification of JAD as an eastern dialect, YAD can also be classified as an eastern Arabic dialect due to the comparable values which resulted from the consonantal and vocalic durational measurements of both dialects.

However, as mentioned earlier, the homogeneity of the participants in terms of their gender and educational background could be a possible factor contributing to the similarity of their rhythm output. It might therefore be expected that by investigating the rhythmic structures of speakers of different educational backgrounds might yield different results. The fact that the study was conducted in a non-native context where there are more chances for speakers to minimize the use of their local dialects can also be a factor that affected the findings of the study. Thus these dialects need to be further investigated in later research in native contexts. Besides, further comparisons need to be made with more Arabic dialects, particularly those of the west, to investigate whether or not significant differences do occur between the eastern and western Arabic dialect.

In conclusion, although research on speech rhythm based on the stress timing and syllable timing features of languages is no longer a young field since it dates back to the 1940s, it is still widely investigated by researchers to compare between languages and capture the cross linguistic rhythmic variations. Researchers have come to know that rhythm is a result of some phonetic properties caused by the prominence or elision of some parts of speech that can be captured acoustically or recognized perceptually as a kind of pattern. However, the concept of rhythm have moved long way from its perception in terms of dichotomous nature of rhythm to a bi-polar continuum into what Nolan and Asu (2009, p. 64) describe as orthogonal dimensions of the rhythmic structure of languages where stress timing and syllable timing features are not necessarily at opposite ends of a continuum since some languages may display both kinds of rhythmic structures.

5.3 Implications of the Study

The results of the current study revealed that the consonantal and vocalic durations can lead to different classifications of the language. Like Asu and Nolan (2006) who suggested that languages can be described by means of a two-dimensional characterization, this study also implies that the two parameters of the consonantal and vocalic intervals provide a better reflection of the rhythm of YAD and JAD. In having different classifications of these Arabic dialects based on the consonantal and vocalic intervals, the concept of coexisting rhythm adopted from Nolan and Asu (2009, p. 76) has been enhanced in this study as a more appropriate notion for dealing with cases like YAD and JAD where two different rhythmic structures appeared to exist within the same language.

In comparing YAD rhythm to that of JAD, no significant differences have been found in the consonantal and vocalic durations of both dialects. JAD has always been classified by researchers and scholars as being an eastern dialect, thus, given the comparable rhythmic structures of both YAD and JAD in this study, YAD can also be classified as an eastern dialect.

5.4 Limitations of the Study

The current study has investigated the acoustic correlates of the rhythm of YAD and JAD by analyzing the consonantal and vocalic durational intervals of the two Arabic dialects. However, the study counted on the use of the PVI only and other indexes were not applied in the study. Besides, the durations of the consonantal and vocalic intervals were taken as two parameters for the description of YAD and JAD but other parameters like the syllable and foot were not examined in the study. The study also applied an acoustic investigation of rhythm without a perceptual analysis of the speakers' rhythm output.

In describing Dauer's theory as a source from which later studies have been inspired to quantify her ideas by proposing different rhythmic indexes, Dauer's criteria (namely, syllable structure, vowel reduction and lexical stress) have been presented in this study as a theoretical framework. However, only the syllable structure was examined in this study due to time constraints.

5.5 Recommendation for Further Research

The current study is an exploratory study that attempts to identify the rhythm of YAD and JAD and it can be considered as the first of its kind since YAD rhythm has not yet been investigated as far as known by the researcher. However, it targeted speakers in a foreign context where their native dialects are rarely used and heard. Thus, a replicated study that targets speakers in their native context is highly useful and might yield different findings. It is also recommended to choose speakers from different backgrounds to see if these differences might produce different results and yield significant differences between their values. Moreover, YAD was compared to one eastern Arabic dialect only, therefore, it is recommended to include in further research a wider range of Arabic dialects from the east and west to achieve a thorough investigation and then further comparisons can be made with other languages of the world.

This research has investigated the YAD and JAD rhythm based on two rhythmic parameters, namely, the consonantal and vocalic durations that have been analyzed by the *rPVI* and *nPVI* respectively, yielding two different classifications of these Arabic dialects. It is recommended to further investigate the YAD and JAD rhythm by analyzing the foot and syllable structure and by applying more indexes for the sake of validation and comparison between different indexes. It should also be noted that this study measured one aspect of rhythm, which is timing; thus, other properties of speech can be considered for investigating rhythm in future research. Besides, although this study focused only on rhythm, it can also pave the path for investigating other suprasegmental features of YAD and JAD like stress and intonation.

REFERENCES

- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Adams, C. (1979). *English speech rhythm and the foreign learner*. The Hague: Mouton.
- Al-Ani, S. (1970). *Arabic phonology: An acoustical and physiological investigation*. The Hague/Paris: Mouton.
- Alias, A. (2003). Style stratification in the phonology of Malaysian learners of English. In Hajar Abdul Rahim and Shakial Abdul Manan (Eds.), *Linguistics and English language teaching*. Malaysia: Penerbit Universiti Sains Malaysia.
- Allen, G. D., & Hawkins, S. (1979). Trochaic rhythm in children's speech. In H. Hollien & P. Hollien (Eds.), *Current issues in the phonetic sciences* (pp. 927-933). Amsterdam: John Benjamins.
- Allen, G. D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G.H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), *Child phonology: Production* (pp. 227-256). New York: Academic Press.
- Al-Masri, M., & Jongman, A. (2004). Proceedings of the 2003 Texas Linguistic Society Conference: *Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results*. Smorville, Massachusetts Cascadilla Proceedings Project.
- Al-Saidat, E., M. (2010). Phonological analysis of English phonotactics: A case study of Arab learners of English. *The Buckingham Journal of Language and Linguistics*, 3, 121-134.
- Al-Shuaibi, A. (2009). Phonological analysis of English phonotactics of syllable initial and final consonant clusters by Yemeni speakers of English. *Language in India*, 9(11), 195-328
- Al-Yaari, S. A., Al-Hammadi, F. S., & Luwa, A. B. (2012). Vowel deletion in Arabic dialects of Yemen: A linguistic perspective. *International Journal of Phonetics*, 2(2), 430-458.
- Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, (66) 46-63.
- Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, (40) 351-373.
- Asu, E.L., & Nolan, F. (2006). Proceedings of Speech Prosody: *Estonian and English rhythm: A two-dimensional quantification based on syllables and feet*. TUD press, Dresden.

- Awde, N. & Samano, P. (1986). *The Arabic alphabet: How to read & write*. London: British library.
- Babbie, E. (2005). *The basics of social research* (3rd ed). Canada: Chapman University
- Balasubramanian, T. (1980). Timing in Tamil. *Journal of Phonetics*, (8) 449-467.
- Ball, M.J., & Rahilly, J. (1999). *Phonetics: The science of speech*. London: Arnold.
- Barkas, P. (1934). A critique of modern English prosody. *Studien zur Englischen philologie (Halle)*, 82.
- Barry, W. J., & Andreeva, B. (2001). Cross-language similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association*, 31(1), 51-66.
- Barry, W. J., Andreeva, B., Russo, M., Dimitrova, S., & Kostadinova, T. (2003). Proceedings of the 15th International Congress of Phonetic Sciences: *Do rhythm measures tell us anything about language type?*, Barcelona.
- Bedawi, E. M. (1973). *Mustawayāt al-'arabiyya al-mu'aşira fī mişr*. Cairo: Dār al-Ma'ārif.
- Benton, M., Dockendorf, L., Jin, W., Liu, Y., & Edmonson, J.A. (2007). Proceedings of the 16th International Congress of Phonetic Sciences: *The continuum of speech rhythm: Computational testing of speech rhythm of large corpora from natural Chinese and English speech*. Saarbrücken (Germany).
- Bertinetto, P. M., & Bertini, C. (2008). Paper presented at Speech Prosody: *On modeling the rhythm of natural languages*. Campinas, Brazil.
- Blanc, H. (1960). Style variation in spoken Arabic: A sample of interdialectal conversation. *Contributions to Arabic linguistics*. Cambridge: Harvard University Press.
- Bloch, B. (1950). Studies in colloquial Japanese IV: Phonemics. *Language*, 26, 86-125.
- Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.17, 2012. [Online] Available: <http://www.praat.org/>.
- Bolinger, D. (1965). *Pitch accent and sentence rhythm in forms of English: Accent, morpheme, order*. Harvard: Harvard University Press.
- Brown, W. (1911). Temporal and accentual rhythm. *Psychological Review*, 18(5), 336-346.
- Bunta F., & Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish- and English-speaking 4- and 5-year-old children. *Journal of Speech, Language, and Hearing Research*, 50(4), 999-1014.

- Chapman, J. (1818). *The music or melody and rhythmus of language*. Edinburgh: Anderson.
- Chatman, S. (1965). *A theory of meter*. The Hague: Mouton.
- Clark, J., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology* (3rd ed). United Kingdom: Blackwell Publishing.
- Classe, A. (1939). *The rhythm of English prose*. Oxford: Blackwell.
- Collins English Dictionary* (1st ed). Glasgow: HarperCollins.
- Crystal, D. (1985). *Dictionary of linguistics and phonetics* (2nd ed.). Oxford: Blackwell.
- Crystal, D. (1990). *A dictionary of linguistics and phonetics* (3rd Edition). London: Blackwell.
- Crystal, D. (2003). *English as a global language* (2nd ed). Cambridge: Cambridge University Press.
- Crystal, D. (2007). *The Cambridge encyclopedia of Language* (2nd ed). Cambridge: Cambridge University Press.
- Cumming, R. E. (2010). *Speech rhythm: The language-specific integration of pitch and duration* (Doctoral dissertation). Downing College, Cambridge.
- Dasher, R., & Bolinger, D. (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association*, 12, 58-69.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- Dauer, R. M. (1987). Paper presented at the 11th International Congress of Phonetic Sciences: *Phonetic and phonological components of language rhythm*. Tallinn, Estonia.
- Dellwo, V. (2006). Language and language processing: Proceedings of the 38 linguistic colloquium: *Rhythm and speech rate: A variation coefficient for delta C*. P. Karnowski & I. Szigeti (Eds.). Frankfurt: Peter Lang.
- Dellwo, V., & Wagner, P. (2003). Proceedings of the 15th International Congress of Phonetics Sciences: *Relations between language rhythm and speech rate*. Barcelona (Spain).
- Deterding, D. (1994). Paper presented at the 5th Australian International Conference on Speech Science and Technology: *The rhythm of Singapore English*. Western Australia.

- Deterding, D. (2001). The measurement of rhythm: A comparison of Singapore English and British English. *Journal of Phonetics*, 29(2), 217–230.
- Deterding, D. (2011). Proceedings of ICPS: *Measurements of the rhythm of Malay*.
- Dickins, J. (1998). *Extended axiomatic linguistics*.: Berlin and New York: Mouton de Gruyter.
- Draat, P.F.V. (1909). *Rhythm in English prose*, Heidelberg: Winter.
- Duckworth, J. (1967). In D. C. Wigglesworth (Eds.), selected conference papers of the association of teachers of English as a second language: *The rhythm of spoken English*. Los Altos, California: Language Research Associates' Press.
- El-Hassan, S. (1994). English accentuation and vowel quality as pronounced by Arabs : A pedagogic statement. *Papers and Studies in Contrastive Linguistics*. (29), 205-216.
- El-Isa, S. (1982). *A study of dynamic features in Irtah spoken Arabic with particular reference to stress, rhythm and intonation* (Doctoral dissertation). University of Glasgow.
- El-Medlaoui, M. (1995). *Aspects des représentations phonologiques dans certaines langues camitosemitiques*. Publication de la Faculte des Lettres et des Sciences Humaines de Rabat.
- Feagin, Crawford. (2008). Entering the community: Fieldwork. In Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (Eds.), *The Handbook of Language variation and Change* (pp. 20-39). Oxford: Blackwell publishers.
- Fischer, W. (1997). *Classical Arabic: The Semitic Languages*. London: Routledge.
- Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford: Oxford University Press.
- Frota, S., & Vigario, M. (2001). On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus*, 13(2), 247- 275.
- Fry, D.B. (1964). The function of the syllable. *Zeitschrift fur Phonetik*, (17)215-221.
- Fujisaki, H. (1997). In, Y. Sakisaka, Nick Campbell & N. Higuchi (Eds.), Computing prosody: *Prosody Models and Spontaneous Speech*. New York: Springer.
- Ghazali, S. (1977). *Back consonants and backing coarticulation in Arabic* (Doctoral dissertation). University of Texas.
- Ghazali, S., Hamdi, R. & Barkat, M. (2002). Proceedings from Speech Prosody: *Speech rhythm variation in Arabic dialects*. Aix-en-Provence.

- Ghazali, S., Hamdi, 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). Paper presented at Eurospeech: *Measuring speech rhythm*. Aalborg, Denmark.
- Grabe, E., Gut, U., Post, B., & Watson, I. M. C. (1999). *The Acquisition of Rhythm in English, French and German*. Paper presented at The Child Language Seminar, London, UK 1999.
- Grabe, E., & Low, E. L. (2002). In N. Warner & C. Gussenhoven (Eds.), *Papers in laboratory phonology VII: Durational variability in speech and the Rhythm Class Hypothesis*. Berlin: Mouton de Gruyter.
- Guest, E. (1838). *History of English rhythm*. London: George Bell & Sons.
- Gut, U. (2012). Rhythm in L2 speech. *Speech and Language Technology*, 14(15), 83–94.
- Hamdani, G. D., Selouani, S. A., & Boudraa, M. (2010). Algerian Arabic speech database: Corpus design and automatic speech recognition application. *Arabian Journal For Science and Engineering*, 35(2c), 157-166.
- Hamdi, R., Barkat, M., Ferragne, E., & Pellegrino, F. (2004). Proceedings of Interspeech: *Speech timing and rhythmic structure in Arabic dialects: A comparison of two approaches*. Jeju, Korea.
- Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.
- Harris, J. M., & Gries, S. (2011). Measures of speech rhythm and the role of corpus-based word frequency: A multicultural comparison of Spanish-English speakers. *International Journal of English Studies*, 11(2), 1-22.
- Harry, B. (1996). The importance of the language continuum in Arabic multi-glossia. In A. Elgibali (Eds.). *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi* (pp.66-90). Cairo: The American University in Cairo Press.
- Hassan, T. (1955). *Linguistic methods in language research*. Cairo: Anglo-Egyptian Bookshop.
- Heliel, M.H. (1977) *The rhythm of Egyptian colloquial Arabic: An experimental study*. (Doctoral dissertation). London University College.
- Herries, J. (1773). *The elements of speech*. London: Edward and Charles Dilly.

- Hoequist, C. (1983b). Syllable duration in stress-, syllable- and mora-timed languages. *Phonetica*, 40, 203-237.
- Jalal, A. (2011). Pronunciation problems among Saudi learners: A case study at the preparatory year program, Najran University, Saudi Arabia. *Language in India*, 11(7), 22-36.
- Jian, H. (2004). Proceedings of Speech Prosody: *On the syllable timing in Taiwan English*. Nara (Japan).
- Jones, D. (1960). *An outline of English phonetics* (9th ed). Cambridge: Heffer.
- Keller, E., & Zellner, B. (2000). The chaotic nature of speech rhythm: Hints for fluency in the language acquisition process. In Delcloque, Ph., Holland, V.M. (Eds.), *Integrating speech technology in language learning*. London: Swets & Zeitlinger.
- Kent, R. D., & Read, C. (2002). *Acoustic analysis of speech* (2nd ed). USA: Delmar Learning.
- Kingdom of Jordan*. Retrieved January 2014, from Wikipedia website, <http://en.wikipedia.org/wiki/Jordan#Languages>
- Labov, W. (1966). *The social stratification of English in New York city*. Washington, DC: Center for Applied Linguistics.
- Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladefoged, P. (1973). The features of the larynx. *Journal of Phonetics*, 1, 73-84.
- Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sounds of languages*. USA: Blackwell Publishers.
- Ladefoged, P. (2006). *A course in phonetics* (5th ed). Los Angeles: Michael Rosenberg.
- Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th ed). Canada: Wadsworth Learning.
- Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.
- Lin, H., & Wang, Q. (2008). Mandarin rhythm: An acoustic study. *Journal of Chinese Language and Computing*, 17(3), 127-140.
- Lloyd James, A. (1940). *Speech signals in telephony*. London: Sir Isaac Pitman & Sons.
- Loukina, A., Kochanski, G., Shih, C., Keane, E., & Watson, I. (2009). Proceedings of Interspeech: *Rhythm measures with language-independent segmentation*.

- Low, E. L. & Grabe, E. (1995). Proceedings of the International Congress of Phonetic Sciences: *Prosodic patterns in Singapore English*. , Stockholm.
- 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.
- Lucas, D. W. (1968). *Aristotle's poetics*. Oxford: Clarendon Press.
- Mairano, P. (2010). *Rhythm typology: Acoustic and perceptive studies* (Doctoral dissertation). Turin University, Italy.
- Meiseles, G. (1980). Educated spoken Arabic and the Arabic language continuum. *Archivum Linguisticum*, 15(2), 118-143.
- Miller, J. G. (1978). *Living systems*. New York: McGraw Hill.
- Miller, M. (1984). On the perception of rhythm. *Journal of Phonetics*, 19, 231-248.
- Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and interpretation*. Oxford: Blackwell.
- Mok, P. K., & V. Dellwo (2008). Proceedings of Speech Prosody: *Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English*. Campinas, Brazil.
- Na'ama, Abbas. (2011). An analysis of errors made by Yemeni university students in the English consonant clusters system. *Damascus University Journal*, 27(3+4), 145-161.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756-766.
- Nespor, M. (1990). *On the rhythm parameter in phonology. The logical Problem of language acquisition*. Dordrecht: Foris
- Newman, D., & Verhoeven, J. (2002). Frequency analysis of Arabic vowels in connected speech. *Antwerp Papers in Linguistics*, 100, 77-86.
- Nakatani, L.H., O'Connor, K.D., & Aston, C.H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica*, 38, 84-106.
- Nolan, F. & Asu, E. L. (2009). The Pairwise Variability Index and coexisting rhythms in language. *Phonetica*, 66(1-2), 64-77.
- O'Connor, J. D. (1965). The perception of time intervals. *Progress Report, Department of Phonetics, University College London*, 2, 11-15.

- Omond, T.S. (1921). *English metrists*. Oxford: Oxford University Press.
- Oxford English dictionary*. (1989). (2ed.) Oxford: Oxford University Press.
- Patmore, C. (1857). English metrical critics. *North British Review*, 27(53), 127-61.
- Pointon, G. (1980). Is Spanish really syllable-timed? *Journal of Phonetics*, 8, 293-304.
- Port, R., Dalby, J., & O'Dell, M. (1987). Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America*, 81(5), 1574-1585.
- Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Rammuny, R. (1966). *An analysis of the difference in the prosodies of general American English and colloquial Jordanian Arabic and their effect on second language acquisition*. Ann Arbor: CRLLB, The University of Michigan.
- Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- Republic of Yemen* . Retrieved January 2014, from Wikipedia website, http://en.wikipedia.org/wiki/Yemen#Administrative_divisions
- Resto, J. (2000). Proceedings of the Seminar for Arabian Studies: *When and what was Arabic Felix?*, 30, 189-192.
- Schmid, S. (2001). *Un nouveau fondement phonétique pour la typologie rythmique des langues?* Poster presented at the workshop for 'le 10ème anniversaire du Laboratoire d'analyse informatique de la parole (LAIP)', Lausanne, Switzerland, 31st May 2001.
- Schmidt, R. W. (1974). *Sociolinguistic variation in spoken Egyptian Arabic: A re-examination of the concept of diglossia* (Doctoral dissertation). Brown University.
- Shar, S., & Ingram, J. (2010). Proceedings of the 13th Australasian International Conference on Speech Science and Technology: *Pharyngealization in Assiri Arabic: An acoustic analysis*. Melbourne, Australia.
- Shen, Y., & Peterson, G. G. (1962). *Studies in Linguistics*: University of Buffalo Occasional Papers: *Isochronism in English*, 9, 1-36.
- Sonnenschein, A.E. (1925). *What is rhythm?* Oxford: Blackwell.

- Sokan almamlaka*. Retrieved January 2014, from Estimated Population of the Kingdom by Governorate website, http://dos.gov.jo/dos_home_e/main/ehsaat/alsokan/2011/
- Soraya, H.A.I. (1966) An intonational study of Egyptian colloquial Arabic (Doctoral dissertation). University of London.
- Steele, J. (1775). *Prosodia rationalis: Or, An essay towards establishing the melody and measure of speech*. London: J. Nichols.
- Sumera, M. (1968). The temporal tradition in the study of verse structure. *Department of Linguistics, Edinburgh University*, 62, 21-43.
- Sweet, H. (1913). Words , logic, and grammar. TPS 1875-6, collected papers of H. Sweet Oxford.
- Tajima, K., Zawaydeh, B. A., & Kitahara, M. (1999). Poster presented at XIVth International Congress of Phonetic Sciences (ICPhS99): *A comparative study of speech rhythm in Arabic, English, and Japanese*. San Francisco, CA.
- Tan, R. S. K.(2011). An acoustic investigation of the segmentals and suprasegmentals of Malaysian English. (Doctoral dissertation).
- Tan, R. S. K., & Low, E. L. (2013). Rhythmic patterning in Malaysian and Singaporean English. *Language and Speech*, 1-19.
- Tannen, D. (1990). *Gender differences in conversational coherence: Physical alignment and topical cohesion*. Norwood, NJ: Ablex Publishing Corp.
- Thelwall, R., & Saadeddin, M. K. (2003). Arabic. In *Handbook of the international phonetic association* (pp. 51-54). Cambridge: Cambridge University Press.
- Thomas, E., & Carter, P. M. (2006). Prosodic rhythm and African American English. *English World Wide*, 27(3), 331–355.
- Thomson, W. (1926). *The rhythm of Greek verse*. Glasgow: Maclehose & Jackson.
- Tortel, A & Hirst, D. (2008). Rhythm and rhythmic variation in British English: Subjective and objective evaluation of French and native speakers. In P.A. Barbosa, S. Madureira & C. Reis (Eds.) *Proceedings of Speech Prosody* (pp. 359-362) Campinas, Brazil: Pablo Arantes
- Trask, R. L. (2007). *Language and linguistics* (2nd ed). NewYork: Routledge.
- Wagner, P., & Dellwo, V. (2004). Paper presented at *Speech Prosody: Introducing YARD (Yet Another Rhythm Determination) and reintroducing isochrony to rhythm research*. Nara, Japan.

- Wallin, J. E. W. (1901). Studies from the Yale Psychological Laboratory: *Researches on the Rhythm of Speech*, 9, 1-142.
- Watson, J. C. E. (2002). *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.
- Wenk, J. W. & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10, 193-216.
- White, L. & Mattys, S. L. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.
- White, L., & Mattys, S. (2007b). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró & M.-J. Solé (Eds.), *Segmental and prosodic issues in romance phonology* (pp. 237-257). Amsterdam: John Benjamins.
- Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. L. (2009). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*, 127, 1559-1569.

APPENDIX A

Read Passage Material

كانت ريحُ الشمال تتجادلُ والشمسُ في أي منهما كانت أقوى من الأخرى وإذ بمسافرٍ يطلعُ متلفعاً بعبائةٍ سميكةٍ. فاتفقتا على اعتبار السابق في اجبار المسافر على خلع عباءته الأقوى. عصفت ريحُ الشمال بأقصى ما استطاعت من قوى ولكن كلما ازداد العصفُ ازدادَ المسافرُ تدثراً بعباءته الى أن أسقط في يد الريح فتخلت عن محاولتها. بعدئذٍ سطعت الشمسُ بدفئها فما كانَ من المسافر الا أن خلعَ عباءته على التو. وهكذا اضطرت ريحُ الشمال الى الاعتراف بأن الشمس كانت هي الأقوى.

APPENDIX B

Translation of the Read Passage Material

The North Wind and the Sun were disputing which was the stronger when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the sun shone up warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the sun was stronger of the two.

APPENDIX C

Background of Yemeni Participants

SPEAKER	SEX	AGE	NATIVE LANGUAGE	EDUCATIONAL LEVEL	COURSE
1	F	34	YAD	PhD	Business Administration
2	F	31	YAD	PhD	Business Administration
3	F	28	YAD	Master	Computer Science
4	F	30	YAD	PhD	Medicine
5	F	26	YAD	Master	Civil Engineering

APPENDIX D

Background of Jordanian Participants

SPEAKER	SEX	AGE	NATIVE LANGUAGE	EDUCATIONAL LEVEL	COURSE
1	F	34	JAD	PhD	Islamic Studies
2	F	28	JAD	Master	Civil Engineering
3	F	30	JAD	PhD	Computer Science
4	F	35	JAD	PhD	Mechanical Engineering
5	F	25	JAD	Master	Computer Science

APPENDIX E

Consent of Participants

Name: _____ Sex: ____ Age: ____
Occupation: _____ Nationality: _____
Address: _____
Tel. No. _____ Email: _____

Relationship with the interviewer: _____

Date of Recording: _____

I agree to allow my recordings to be used for research purposes.

Signature: _____