Speech Rhythm in English and Arabic: A Contrastive Review

Mohammed Nour Abu Guba   Shehdeh Fareh   Sane Yagi
University of Sharjah, UAE

Abstract: This paper compares Arabic and English speech rhythms to increase awareness of this neglected and often misunderstood topic in foreign language acquisition. Unlike previous studies, we adopt a phonological view of speech rhythm rather than an isochrony-based phonetic view. We detail the components of speech rhythm at the word and utterance levels in Arabic and English focusing on the rhythmical differences that would affect the learners' rhythm of both languages negatively. Findings suggest that Modern Standard Arabic (MSA) and Jordanian-Ammani Arabic (JAA), unlike English, should be placed at the lower end of the rhythmic continuum. The study opens new directions for future research and concludes with pedagogical implications for learners of Arabic and English.

1. Introduction
Speech rhythm refers to the systematic temporal organization of prominent and less prominent speech units, which include segments, syllables, feet, words, and phrases (Fletcher 2010). It is the alternation of more and less audible units that establishes the speech rhythm of a given language (Nespor, Shukla & Mehler 2011). It is well-established in the literature that English has a stress-timed rhythm with an alternation of stressed and unstressed units in an utterance where stressed units tend to recur at regular intervals (Abercrombie 1967; Dauer 1983; Ramus, Nespor & Mehler 1999; Ladefoged & Johnson 2015). Unlike the well-established English rhythm, Arabic rhythm has not received adequate attention, and it is not yet established what exactly alternates in an Arabic utterance (cf. Section 5.3).

Rhythm is very important in spoken language both in first (L1) and second (L2) language acquisition for many reasons. First, rhythm, along with intonation, is the first to acquire by infants (Prieto & Esteve-Gibert 2018). Second, the realization of intonation itself is based on rhythm (Todaka 1990, cited in Celce-Murcia, Brinton, & Goodwin 2010); therefore, without acquiring the right rhythm, it will not be possible to acquire the right intonation. In foreign language acquisition, rhythm also plays a major role. Many studies have reported that suprasegmental aspects (stress, rhythm, and intonation) are far more important than segmental ones as they contribute considerably to intelligibility, comprehensibility, and accentednees (Adams 1979; Dalton & Seidlhofer 1994; Munro & Derwing 1995; Celce-Murcia et al. 2010). Speaking without the right rhythm affects intelligibility negatively (Halliday 1989) and places more cognitive demands on listeners (Southwood & Fledge 1999).

More evidence for the significance of rhythm in L2 acquisition comes from native speakers' (NSs) reactions to rhythmical/irrhythmical speech. Celce-Murcia et al. (2010) reported that NSs were less tolerant of suprasegmental errors than segmental ones, and they went a step further by describing irrhythmical speech as choppy, markedly non-native, and even aggressive.

Rhythm can also help learners perfect their pronunciation of English. By learning that English rhythm requires an alternation of prominent and less prominent units, learners can recognize whether a vowel is to be realized as a full or reduced one, an aspect of English pronunciation that causes difficulty to most learners of English from many backgrounds (Gut 2009; Celce-Murcia et al. 2010; Abu Guba, Mashaqba, Jarbou & Hajeid, in press, a).

Another important reason that makes understanding the mechanisms of English rhythm indispensable relates to understanding the uninhibited pronunciation of English NSs (cf. Celce-Murcia et al. 2010). Non-native speakers (NNSs) have the impression that English NSs speak very fast and do not articulate all the sounds rendering their speech hard to understand. Studying English rhythm will make students aware of the systematic modifications that NSs make to realize the typical English rhythm. Understanding speech rhythm is also crucial in speech pathologies where a shift to syllable-timing is common among autistic or schizophrenic English NSs (Cummins 2015). Also, it is crucial to identify the segmentation unit in speech processing in languages with different rhythms. Research shows that NSs of English base their segmentation on stressed syllables, while NSs of syllable-timed languages (e.g., French) segment at each syllable (Cutler, Mehler, Norris & Segui 1986). Studying Arabic rhythm could shed light on the way Arab speakers process speech.

Despite the importance of rhythm in speech and the great difficulty it causes to foreign/second language learners (Barry 2007; Gut 2009), few studies have dealt with rhythm in Arabic or in the interlanguage (IL) of Arab learners of English). Previous studies focused on the acquisition of segmental aspects (e.g., Farrah & Halahlah 2020); a few tackled suprasegmentals focusing on stress and intonation (e.g., Kharma & Hajjaj 1997), and very few addressed rhythm in Arabic (cf. Section 5). Therefore, there is a dire need to do more research to better understand this neglected area. This paper will conduct a comprehensive study that compares Arabic and English rhythms from a phonological perspective rather than an isochrony-based phonetic perspective (see Section 2).

By comparing Arabic and English rhythms, the current paper will build awareness of the nature and importance of speech rhythm and stimulate interest and pave the way for future research on rhythm in Arabic and in the IL of Arab learners of English. This study will thus provide a framework for future research that will eventually enable us to better understand Arabic rhythm characteristics and their influence on the acquisition of English rhythm. This study will also contribute to the contrastive analysis literature on Arabic and English --a comparison that is not only essential for Arab learners of English but also learners of Arabic, especially English ones. More specifically, the current paper attempts to answer the following questions:

1. What are the components of speech rhythm in English and Arabic?
2. What are the similarities and differences between Arabic and English speech rhythms?
3. Is Arabic rhythm stress-timed or syllable-timed?

By answering the first two questions, we will highlight the difficulties that L2 learners may encounter in acquiring speech rhythm. Although English learners of Arabic will also benefit from such a comparison, our focus will be on the problems facing Arab learners.

To answer the research questions above, the researchers adopted a contrastive analysis approach where similarities and differences between Arabic and English rhythms were identified. Data on English came from already published literature. However, much of the data on Arabic came from a large ongoing project (led by the first researcher) that acoustically investigates speech rhythms in MSA and JAA. (Note that JA will be used when reference is made to Jordanian Arabic in general). Some data came from earlier research, whenever available. When no data was available on a certain topic, the researchers conducted pilot studies based on the pronunciation of 10 female JAA monolingual speakers whose ages ranged from 18 to 26 (mean=22). They lived in Amman and had never resided outside Amman for more than a month. All studied in Arabic-medium governmental schools, and their knowledge of English is limited. Six speakers had a high school diploma, and the rest had a two-year college degree in Education, Arabic, or Islamic studies. The participants were approached by the first researcher and were requested to read/pronounce words/sentences in JAA (details are provided where appropriate in the sections below). Recordings were done using an LG professional recorder at a 44k sampling rate. Acoustic measurements were performed using Praat 1.4.9 (Boersma & Weenink 2015).

In the next section, we provide background information on speech rhythm and its classification, and we then compare Arabic and English rhythm components at the word and the sentence/utterance levels.

2. Background
In this section, we give background information on the classification of rhythm. We show that the isochrony view cannot account for the differences between languages and a phonological view that takes the different phonetic and phonological properties of languages into consideration is better able to classify speech rhythm.

2.1. Classification of rhythm
The study of speech rhythm has attracted the attention of many researchers over the past 100 years (e.g., Abercrombie 1967; Roach 1982; Dauer 1983; Arvaniti 2012). Languages are classified into three classes: stress-timed, syllable-timed, and mora-timed (Pike 1945; Abercrombie 1967; Ladefoged & Johnson 2015). In stress-timed languages such as English and Dutch, the interstress intervals (feet) are isochronous, whereas, in syllable-timed languages such as Spanish and French, the syllables are equally spaced, and in mora-timed languages such as Japanese, moras

are isochronouns (Rammuny 1989; Ladefoged & Johnson 2015). This view of rhythm is known as the 'isochrony view'.

In stress-timed languages, timing depends on the number of stresses, which tend to come at regular intervals (Rammuny 1989; Colantoni, Steele & Escudero 2015; Ladefoged & Johnson 2015). This leads to irregular syllable duration depending on the number of syllables intervening between stressed syllables. Put differently, speakers would compress or stretch syllables to make them fit into the interstress intervals (the typical foot duration). However, in syllable-timed languages, timing depends on the number of syllables whether stressed or not; syllables tend to come at recurrent intervals, with similar durations that are commensurate with their structural makeup, regardless of their stress status (Ladefoged & Johnson 2015).

Studies that examined this isochrony view did not find solid empirical evidence to support it; feet in stress-timed languages were not isochronous; rather, foot duration in stress-timed and syllable-timed languages was found to be proportional to the number and types of syllables (Roach 1982; Dauer 1983; Fletcher 2010; Arvaniti 2012). Such findings called the isochrony view into question. However, perceptual studies of rhythm suggest that languages of different speech rhythms sound different (Munro 1995; Nazzi, Bertoncini & Mehler 1998; Fletcher 2010). Several observations provide evidence for the psychological reality of rhythm. First, people tend to hear non-isochronous speech as isochronous, i.e., they impose isochrony on speech (Fletcher 2010). Similarly, babies and adults can differentiate between languages of different rhythms, but they cannot distinguish between languages of similar rhythms (e.g., Nazzi et al. 1998). Again, in second language speech, it was possible to identify non-native accents based on filtered speech, i.e., prosody only (Munro 1995). This means that rhythm is psychologically real, so where does this perceptual isochrony come from?

Dauer (1983, 1987) suggests that stress/syllable timing results from the different phonological properties of languages, not from isochrony. Dauer (1983:54) found that interstress intervals in English, a prototypical stress-timed language, were no more isochronous than those in Spanish, a syllable-timed language. The most important features that affect rhythm relate to syllable structure, vowel reduction, and stress (Dauer 1987, see Section 3). Dauer (1987) further argued that the interaction of these phonological and phonetic properties (which also affect the duration of consonantal and vocalic intervals) in a certain language makes the language more, or less stress-timed on a continuum of stress-timing.

One problem with this view is that it cannot classify languages based on these properties (Ramus et al. 1999). How much does each property contribute to the speech rhythm? To solve this problem, Ramus et al. (1999) suggested that these properties have acoustic correlates that can be measured reliably. This gave birth to 'rhythm metrics'-- mathematical formulas that compare the duration of consonantal and vocalic intervals. The most common metrics that have been successfully implemented in the classification of rhythm are %V (the proportion of the vocalic intervals in an utterance), ΔV/C (the standard deviation of vocalic/consonantal intervals), and the pairwise variability indices (the degree of contrast between

successive durations of vowels/consonants (Arvaniti 2009, 2012). Stress-timed languages are believed to have lower %V but higher delta-C and delta-V due to the complexity and variation of their syllable structure and vowel reduction, while syllable-timed languages tend to have higher %V but lower ΔV/C (Ramus et al. 1999). For the pairwise variability index, the higher the index is, the more stress-timed a language is.

Studies employing such metrics on Arabic rhythm are rather scarce. The few studies that used these metrics (e.g., Hamdi, Barkat, Ferragne & Pellegrino 2004; Ghazali, Hamdi & Knis 2007) were based on a small number of speakers and their methodologies were not without criticism. Hence, more large-scale studies with robust methodology are needed.

Having established the classification of speech rhythm, in the following section we compare English, a prototypical stress-timed language, with MSA and JAA. This comparison will shed light on the controversial classification of Arabic rhythm.

3. Components of rhythm at the word level

In this section, we compare the most important properties of English, MSA, and JAA, focusing on the three most influential properties, namely syllable structure, stress, and vowel reduction. There is a consensus that English is a prototypical stress-timed language (cf. Section 1); therefore, comparing English and Arabic properties will shed light on the controversial classification of Arabic rhythm (cf. Section 5.3); if Arabic and English share many properties, then it would be reasonable to classify Arabic as stress-timed; otherwise, it would be syllable-timed, or at least, not as stress-timed as English is.

3.1. Syllable structure

Phonologically, a syllable comprises an onset (a consonant or more), a nucleus (usually a vowel sound), and a coda (a consonant or more closing the syllable) (Roach 2000). Unlike MSA and JA, English has a very complex syllable structure with 20 syllable types (Harris 1994). The onset ranges from zero to three consonants, as in 'at' and 'strong', respectively; the nucleus can be a short vowel, long vowel, diphthong, or even triphthong, as in 'cut', 'caught', 'coat', and (British English) 'hour', respectively; and the coda ranges from zero to four, as in 'to', 'ten', 'tens', 'tenths', and 'texts', respectively (Harris 1994; Roach 2000).

On the other hand, JAA and MSA have a simpler syllable structure. MSA has only six syllable types (Al-Ani & May 1973; Khalifa 2017), while JAA has 11 (Abu Guba 2016). In MSA, the onset is obligatory, but complex onsets are not allowed. The nucleus can be short, long, or a diphthong, while the coda is usually simplex unless in pause form where it can be composed of two consonants (e.g., *ki.taab* 'book', *nawm* 'sleep', and *nahr* 'river'). JAA has a more complex syllable structure than MSA but still simpler than that of English. Like MSA, an onset is obligatory. The norm is a simplex onset, but two-consonant onsets arise from vowel syncope, as in /bila:d/> [bla:d] 'countries' where the short high vowel /i/ is deleted in unstressed open non-final syllables. The nucleus in JAA is like that in MSA. Again, the optimal coda in JAA is simplex, whereas two-consonant codas are optional and

attested on suffixation or in word-final position, e.g., [kalb ~kalib] 'dog', but [nisir] 'eagle' (Abu Guba 2016).

 Comparing the phonotactics of Arabic and English, we notice that the most noticeable differences between them relate to the onset and coda. The onset in Arabic is obligatory but optional in English, which leads many Arab learners to insert an onset (usually the glottal stop) to English onsetless syllables (Kharma & Hajjaj 1997). Also, English complex onsets and codas are usually realized with an epenthetic vowel in the IL of many Arab learners, as in 'screen'>*[sikriːn] and 'asked'>*[ɑːskid].

 Note that syllabic consonants are very common in English, but they do not exist in MSA. However, they are attested in JAA with different distributions. A syllabic consonant in JAA obtains word-initially following vowel syncope only if the first consonant is more sonorous than the second, as in [nsuur] 'eagles', but not in [blaad] (cf. Angoujard 1990). This difference in distribution is problematic to Arab learners, who tend to insert a vowel to realize an English syllabic consonant, which sounds odd to English NSs (Celec-Murcia et al. 2010). Such repair processes affect the rhythm of the IL of Arab learners by making it sound less stress-timed. The syllable structure produced is less complex and this affects the percentage of vocalic intervals.

3.2. Stress

Stress refers to prominence whereby a stressed syllable stands out from other syllables for being longer, louder, and higher in pitch (Roach 2000; Celce-Murcia et al. 2010). Arabic and English stress systems differ in three main aspects: phonemicity, assignment rules, and realization.

 Unlike Arabic stress, English stress is phonemic, as in the noun 'ˈimport' and the verb 'imˈport'. In neither MSA nor JAA is stress phonemic (notwithstanding few isolated cases in JA, as in ˈfihimna 'our understanding' vs. fiˈhimna 'he understood us'). In fact, the non-phonemicity of stress in Arabic leads to what is called 'stress-deafness' whereby Arab speakers exhibit an inability to perceive stress correctly even in their native language (Al-Jarrah 2002; Peperkamp & Dupoux 2002).

 Stress assignment in MSA and JAA is fully predictable. Stress is assigned to one of the last three syllables in a word in JAA and almost all words in MSA. If the last syllable is superheavy (i.e., CVVC or CVCC), it receives stress, as in mafaˈtiiħ 'keys'; otherwise, the penult receives the stress if it is heavy (CVV or CVC), as in banˈdoora 'tomatoes'; or else stress falls on the antepenultimate syllable (third from last) regardless of weight, as in ˈmadrasa 'school' and ˈbaladu 'his country' (Hayes 1995; Watson 2011; see Abu Guba 2018 for details). By contrast, English is a free stressed language, i.e., stress can fall on any syllable (Hayes 1995).

 Another important difference between Arabic and English stress systems relates to footing initial syllables. JAA requires initial syllables to be parsed into feet in underived words, while English does not (Abu Guba 2021). All underived words in Arabic start with two light syllables, as in katab 'he wrote' (the first two syllables make up one foot) or with a heavy syllable, as in naxla 'palm tree' (the

first heavy syllable constitutes a foot by itself). This difference predicts that JAA learners will have difficulty pronouncing polysyllabic English words that start with unstressed syllables, such as 'moˈnotonous' and 'moˈnopoly', which are often pronounced *'monoˈtonous' and *'monoˈpoly' by many Arab learners. Worse still is that many Arab learners would produce such light syllables with a geminate, as in 'collect' and 'select' (Abu Guba, under review).

Regarding stress realization, although Arabic and English use the three most important correlates to cue stress (fundamental frequency, intensity, and duration), they differ in the extent to which they use these correlates, particularly duration (de Jong & Zawaydeh 1999; Almbark, Bouchhious & Hellmuth 2014). The duration of stressed vowels in English is considerably longer than that in their unstressed counterparts. Greenberg, Carvey, Hitchcock, and Chang (2003) reported that stressed vowels were 60-100% longer than their unstressed counterparts. By contrast, the difference between stressed vowels and their unstressed counterparts in Arabic is usually less than 10%. De Jong and Zawaydeh (1999) reported that Arabic stressed vowels were about 7.5% longer than unstressed ones (215 vs. 200 ms). On the basis of the production of 10 near minimal pairs by 10 JAA female speakers in a pilot study (cf. Section 1), we found that the average duration of stressed short vowels was 59 ms, while that of unstressed vowels came to 56 ms. This is consistent with the fact that languages with phonemic vowel length (like Arabic) tend not to use duration to signal out stress (Gordon 2014).

Another clear difference in stress realization relates to vowel spectral quality (in terms of F1 and F2). English stressed vowels have a peripheral quality, whereas unstressed vowels are usually reduced to schwa. Conversely, it has been reported that JAA speakers do not use vowel quality to cue stress; no conclusive evidence of vowel reduction concerning spectral quality in unstressed syllables was found (e.g., de Jong & Zawaydeh 1999; Abu Guba et al., in press, a). Vowel reduction is taken up in the following subsection.

Two minor differences between English and Arabic stress systems relate to secondary stress and stress shift. While secondary stress is prevalent in English (Roach 2000), the status of secondary stress in Arabic is controversial (Watson 2011; Khalifa 2017, for Egyptian Arabic). Stress shift applies within phrases to avoid 'stress clash', i.e., two adjacent primary stresses. It is very common in English as in 'Heathrow Airport', which is realized as 'ˈHeathrow ˈAirport', with stress shifting from '*throw'* to '*Hea*'. By contrast, stress shift in Arabic does not seem to be as clear as in English, most probably because stressed syllables in Arabic are not as prominent as in English. These two differences do affect rhythm in that they may lead learners to produce syllables without the required alternation of prominence.

3.3. Vowel reduction

Vowel reduction, a key characteristic of English and arguably a fifth cue of stress as it boosts the contrast between stressed and unstressed syllables, refers to the production of unstressed vowels with a centralized quality resulting usually in schwa (Roach 2000; Collins & Mees 2008). Reduced vowels are considerably

shorter than their full vowel counterparts; Crystal and House (1988) found that English monophthongs are 50% shorter in reduced vowels. Qualitatively, reduced vowels have midrange F1 and F2 values (Sluijter & van Heuven 1996).

Unlike English, Arabic has a weaker acoustic contrast between stressed and unstressed vowels. Vowels do not seem to differ significantly in terms of quality or quantity. Some researchers contend that vowel reduction is completely alien to Arabic (e.g., El-Hassan 1994), while others argue that it is negligible (e.g., Zuraiq & Sereno 2007). Some evidence for the lack of vowel reduction in Arabic comes from research on the IL of Arab learners of English. Kharmah and Hajjaj (1997) and Al-Jarrah (2002) observed that the most obvious characteristic of the IL of Arab learners of English is the lack of vowel reduction of unstressed vowels. Abu Guba et al. (in press, a), comparing reduced English vowels as produced by English NSs and JAA learners, found that vowels produced by JAA learners were significantly longer and different in spectral quality than those produced by English NSs.

Having explored the three most influential rhythm components, in the rest of this section, we present other less influential components that can also affect rhythm.

3.4. Vowel systems

Arabic and English differ in the number of vowels and vowel distribution across the vowel space. English has a larger number of vowels with varied qualities than those in either MSA or JAA. The number of vowels in English (excluding triphthongs) ranges from 15 to 20 depending on the dialect (Roach 2000; Ladefoged & Johnson 2015), whereas the number of vowel phonemes in MSA and JAA is 8 each, including two diphthongs that are usually monophthongized in JAA (Abu Guba 2016; Kalaldeh 2018).

The English vowel system is centripetal, i.e., vowels are distributed across the vowel space and unstressed vowels tend to be realized with a central quality. However, MSA and JAA vowel systems are midway between centripetal and centrifugal (where vowels move away from the center), with JAA being more centripetal than MSA (cf. Odisho 2003). This difference predicts that Arab learners will tend to articulate their vowels with more peripheral qualities, which again will affect their IL rhythm.

3.5. Segmental length contrasts

Segmental length is phonemic for both consonants and vowels in MSA and JAA, but neither consonants nor vowels contrast for length in English. Unexpectedly, JAA resorts to consonant gemination to repair ill-formed prosodic structures as in *duxxaːn* 'smoke' and *gultillu* 'I said to him' (see Abu Guba 2021). Recall that this phenomenon is also attested among JAA English learners, who produce many geminates in their IL, even at very advanced levels (Abu Guba, under review).

3.6. Compensatory/polysyllabic shortening

This phenomenon is very widespread in English whereby a vowel, especially a long vowel, is shortened in polysyllabic words (e.g., Lehiste 1972; Turk & Shattuck-

Hufnagel 2000). Lehiste (1972) found that the initial long vowel /iː/ in 'speed', 'speedy' and 'speediness' becomes shorter the more syllables the word has (266, 150, and 115 ms, respectively). This phenomenon was attributed to a tendency to maintain similar durations of feet in stress-timed languages.

MSA and JAA exhibit polysyllabic shortening, but to a lesser degree than English. Abu Guba, Mashaqba and Huneety (2023) found that the differences in duration between stressed syllables in monosyllabic words and polysyllabic words in MSA were considerably large, with vowels becoming shorter in polysyllabic words. However, unlike in English, there were no differences between vowels in disyllabic and trisyllabic words. This finding suggests that Arabic is not as stress-timed as English is (cf. Section 5.3), and Arab learners will face difficulties applying polysyllabic shortening (Abu Guba, under review).

To sum up, the accumulated differences between English and MSA and JAA with respect to these properties make them sound rhythmically different. All the processes discussed in this section obtain at the word level. In the next two sections, we introduce other phonetic modifications that contribute to the overall speech rhythm at the sentence/utterance level.

4. English rhythm at the sentence/utterance level

The citation forms of English words undergo several modifications to make them fit into the prosodic timing of English rhythm. These modifications result from the requirements of sentence/utterance stress assignment, rhythmical footing, and connected speech processes.

4.1. Sentence/utterance stress

Sentence/utterance stress, together with word stress, shapes the peculiar rhythm of English. Sentence/utterance stress is based on word stress. Within individual English words, rhythmic alternation in polysyllabic words results in a tendency for an alternation of stressed/unstressed syllables. For example, in 'combination' /ˌkɒm.bɪˈneɪ.ʃən/ there are four levels of prominence (secondary stressed, unstressed, primary stressed, and reduced, respectively). Note that the fourth level is not redundant as there is a difference between unstressed and reduced vowels (cf. the first and last syllables in 'translation' /trænsˈleɪ.ʃən/). At the sentence/utterance level, another level is added: the nuclear/phrasal stress, which usually falls on the stressable syllable in the last content word in the unmarked case, as in 'This combination is nice', with five prominence levels with a nuclear stress on 'nice' (Roach 2000). Another example is 'I want to find where he is going to', where nuclear stress falls on the first syllable in 'going', not on 'to' as it is a preposition (a function word). Note that nuclear stress can occur earlier in an utterance for cases of emphasis or contrast, and if there is a word that has greater importance than what is said after it, as in 'Here is the shirt I wanted' with nuclear stress falling on 'shirt' (see Section 4.3 and Roach 2000: Chapter 19 for details).

Besides this tendency for an alternation of prominent and less prominent syllables, English sentence/utterance rhythm is governed by the temporal

organization of sound units such that metrical feet tend to recur at regular intervals of time (Section 4.2) and by connected speech processes (Section 4.3).

4.2. Rhythmical footing

The temporal organization of speech in English depends on the number of stresses rather than the number of syllables. Each primary stress corresponds to a rhythmical foot, which is composed of a stressed syllable and all the syllables that follow it up to but not including the next stress (Roach 2000). To clarify this stress-timed organization of English utterances, consider the following sentences in (1).

     (1) a. (ˈBoys) (ˈride) (ˈbikes).

       b. The (ˈboys have been) (ˈriding the) (ˈbikes).

     In these sentences, there are three stresses, therefore three feet (brackets demarcate feet). These two sentences, though different in the number of syllables (3 vs. 8), tend to take similar time to produce (but are not isochronous). Consequently, the syllable duration will vary greatly. The stressed syllables in (1a) will be longer than in (1b), and the stressed syllables in (1b) will be considerably longer than the unstressed ones. Also, the word 'bikes' in both sentences will be the most prominent syllable as it has the phrasal stress. Note that English has more levels of stress in sentences with morphologically complex words (Celce-Murcia et al. 2010), but practically and pedagogically, we restrict ourselves to 5 levels, namely phrasal/nuclear stress, primary stress, secondary stress, unstressed and reduced.

4.3. Processes in connected speech

Besides the different degrees of stress within an utterance, there are some connected speech processes that affect English utterances. Although these processes tend to be universal (Celce-Murcia et al. 2010), they differ in degree across languages. Utterances in English are organized into thought groups, which roughly correspond to syntactic and semantic units, and these processes usually occur within (but not across) thought groups. The most common processes are contractions and weak forms (e.g., [(ə)n] for 'and'), linking (e.g., inserting [w] in 'you and'> [juwən]), and assimilation (e.g., 'would you'>[wʊdʒjə]). These are natural (not careless or sloppy) processes that help produce smooth, fluent, comprehensible, and even friendly speech in formal and informal speech (Collins & Mees 2008; Celce-Murcia et al. 2010). The faster the speech is, the more processes are used. Failing to use these processes results in a deviant speech rhythm, as observed among many Arab learners who tend to overarticulate their speech (e.g., Rammuni 1989; El-Hassan 1994; Kharma & Hajjaj 1997).

     Before closing this section, consider the example in (2) that exemplifies how stresses are distributed within an English utterance and how processes of connected speech create a smooth transition between words.

(2) Sally and Sam have asked her difficult questions.

This sentence has 5 primary stresses (shown with ˈ) that correspond to five feet that tend to have similar durations, shown between brackets in (3).

     (3) (ˈSally and) (ˈSam have) (ˈasked her) (ˈdifficult) (ˈquestions).

In this sentence, there are 12 syllables with different degrees of stress/prominence. Five syllables have primary stress; 'questions' have the nuclear/phrasal stress; while the other syllables are unstressed or reduced. Other processes of connected speech apply, most important of which are linking (linking /j/ between 'Sally' and 'and', linking word-final consonants to word-initial vowels as in 'have asked'), reduction and deletion (in 'and', 'have', and 'her'), and obligatory assimilation in 'asked', where /d/ assimilates in voicing to the voiceless stop /k/ (the underlined syllable in 4 receives the phrasal accent).

(4) ˈsæli jən ˈsæmə ˈvæsk tər ˈdɪfɪkəlt ˈkwestʃənz.

Hearing such an utterance (with deletions, insertions, and word boundaries blurred) is quite challenging for NNSs. Increased awareness of such characteristics would make NNSs better able to hear English natural speech and ultimately produce it in an English manner. Having explored the mechanisms of English sentence/utterance rhythm, we turn in Section 5 to Arabic sentence/utterance rhythm.

5. Arabic sentence/utterance rhythm

Very few researchers have tackled Arabic sentence/utterance rhythm, which is still an understudied area (Chahal & Hellmuth 2014). Therefore, we relied on the preliminary findings of a pilot study that we conducted when no data was available.

The alternation of prominent/non-prominent units within Arabic utterances is less clear-cut than it is in English. Every word seems to take a degree of stress (Rammouni 1989; Hellmuth 2006), which suggests that Arabic is less stress-timed than English is, and that the number of syllables is more important than the number of stresses. If Arabic were stress-timed as English is, we would expect an utterance to be realized according to the mechanisms of stress-timing. Let us consider the following sentence from JAA as an example.

(5)   ˈhaːj ˈʃadʒara zaˈraʕha saˈʕiːd ˈgabil ˈʃahir. 'This is a tree planted by Saeed a month ago'

(haːj)(ʃadʒaraza)(raʕhasa)(ʕiːd)(gabil)(ʃahir)

If JAA were stress-timed, we would expect the following: stressed syllables would be longer than unstressed syllables, the vowel /a/ in the first syllable in the second foot (ʃadʒaraza) would be longer than the vowels in the other syllables. Feet would not vary significantly in length. Thus, the vowels in the unstressed syllables would undergo shortening, whereas the vowels in monosyllabic feet would undergo lengthening to render feet similar in length.

Based on the pronunciation of 10 JAA female speakers who repeated this sentence from memory (cf. Section 1), a one-way ANOVA test revealed no statistically significant differences in the durations of stressed and unstressed /a/ in *ʃadʒara* (p=0.184). However, there were statistically significant differences between the duration of feet of different lengths (p=.001). In fact, duration was commensurate with the number of syllables in feet. This suggests that stress does not play a major role in syllable duration in Arabic. Stressed syllables are not considerably longer than unstressed ones —an indication that JAA is less stress-timed than English.

5.1. Function words in Arabic

Empirical acoustic studies on function words in Arabic are lacking. Kenworthy (1987) claimed that function words in Arabic (without specifying which Arabic) have one form only and their vowels are like the vowels in content words. His claim is questionable since the citation forms of some function words, (e.g., *fiː 'in'*), are realized differently in connected speech. */fiː/* is realized as [fi] before words beginning with two consonants, as in [fil.bajt] 'in the house', while it is realized as [fiː] before words beginning with one consonant, as in [fiːnaːs] 'there are people'.

Abu Guba, Jarbou, and Abu Qub'a (in press, b) found that MSA and JAA differed from English in that they did not usually make a clear distinction between the vowels of content words and polysyllabic function words, especially in terms of duration. For example, the first vowel in the disyllabic function word *qabla* 'before' was not noticeably shorter than its counterpart in content words. However, vowels in monosyllabic function words were less prominent than their counterparts in content words; they tended to have lower pitch, lower intensity and were usually shorter than their counterparts in content words. For instance, the vowel in the function word *maʕ* 'with' was shorter, less loud, and had lower F0 than its counterpart in content words. Nonetheless, the quality of all types of vowels, as reflected in F1 and F2 frequencies, were similar, an indication that no vowel reduction obtains, unlike in Western Arabic dialects that tend to reduce unstressed vowels (Ghazali et al. 2007). These findings predict that JAA learners would produce English function words with undue prominence. Abu Guba et al. (in press, a) reported that JAA learners failed to verbalize adequately reducible vowels. This is in harmony with Kenworthy's (1987:124) observation that Arab learners have difficulty producing English sentence/utterance stress; they tend to use full forms of function words, which makes them sound as if they were making a contrast or protesting. Ghazali and Bouchhioua (2003) also found that Tunisian Arab learners assign stress to function words, which affects their IL rhythm negatively.

To summarize, Arabic does not seem to exhibit typical temporal characteristics of stress-timed languages: there is no substantial distinction between stressed and unstressed vowels in content words and function words. In the following section, we touch on the most important connected speech processes in MSA and JAA and leave comprehensive large-scale studies for future research.

5.2. Processes of connected speech in Arabic

These processes at the sentence/utterance level affect rhythm in that they change the temporal characteristics of the produced rhythm. Note that some processes also apply at the word level, which is beyond the scope of this paper. The most common processes affecting Arabic utterances include vowel epenthesis, where vowels are inserted to avoid consonant clusters (e.g., /man#lkaːtib/> [manilkaːtib] (MSA) and /miːn#lkaːtib/>[miːnilkaːtib] (JAA) 'who is the writer'); linking, where sounds at word boundaries are linked together (e.g., /fiː#l+madrasa/>[filmadrasa] 'in school'); deletion (e.g., /ʔal+kahf/>[lkahf] in [filkahf] 'in the cave' and /ʔana#allaði/>[ʔanallaði] (MSA) and /ʔana#illi/>[ʔanalli] (JAA) 'I am the one

who'); and assimilation (e.g., /ʔakal+na/>[ʔakanna] 'we ate', /ẖasan#rikib/>[ẖasarrikib] 'Hasan rode', and /miʃ#sahil/>[missahil] 'not easy'.

Note that other properties with a considerable effect on rhythm such as tempo features including speech rate, proportion of pauses, and length of pauses still await thorough studies in Arabic and in the ILs of Arab learners of English.

Now that we have explored the properties that account for the rhythmic distinction between Arabic and English, the question is how Arabic rhythm can be classified?

5.3. Is Arabic stress-timed?

The classification of Arabic rhythm is controversial. Some researchers think that it is stress-timed (Soraya 1966; Abercrombie 1967; Roach 1982; Miller 1984). However, these researchers did not usually specify which Arabic variety they refer to, nor did they support their views with adequate experimental evidence. Rammuny (1989), on the basis of instrumental and auditory experiments, reported that JA is stress-timed but, unlike English, it stresses all words including function words, which makes it 'word-stress timed' rather than phrase-stress timed, as in English. Note that Rammuni mentioned, without reporting any measurements, that the distance between stresses in JA is equal. Tajima, Zawaydeh, and Kitahara (1999) followed Abercrombie but found interstress intervals in JA to be significantly less isochronous than those in English.

Other researchers argue that Arabic is syllable-timed (e.g., Kharma & Hajjaj 1997; Celce-Murcia et al. 2010); however, they did not provide experimental evidence to support their impressionistic views. By contrast, Ghazali, Hamdi, and Barkat (2002), Hamdi et al. (2004), and Ghazali et al. (2007), adopting rhythmic metrics such as %V and ΔC, found that MSA and Eastern Arabic dialects tend to be less stress-timed than Western Arabic dialects. Although their studies were based on a small number of speakers (the number of speakers representing JA or MSA, the focus of this paper, was less than five), and the speakers might not be true representatives of their dialects as they were bilinguals living in foreign countries, the findings of these three studies are important as they show quantitatively that Arabic dialects are rhythmically different. More experimental studies with more speakers to classify Arabic rhythm are highly recommended.

Having established that the isochrony view is untenable (cf. Section 2.1), we draw on the comparisons made in the previous sections between English, a prototypical stress-timed language, and MSA and JAA, and argue that MSA and, to a lesser degree JAA, tend to be more syllable-timed than stress-timed. This contention is consistent with Dauer's (1987) classification where less stress-timed languages tend to have less complex syllable structure, non-phonemic predictable stress mainly cued by pitch contours (rather than duration), and a small degree of vowel reduction.

6. Conclusion and implications

This study has highlighted the most important rhythmical differences between MSA and JAA, on the one hand, and English, on the other. Comparisons suggest that

MSA and JAA are more syllable-timed than stress-timed, and they should be placed toward the low end of Dauer's (1987) rhythmic continuum. Most differences relate to syllable structure, stress, and vowel reduction. MSA and JAA have a simpler syllable structure than that in English. Stress in MSA and JAA, unlike in English, is fully predictable and non-phonemic. Vowel duration does not play a clear role in the distinction between stressed and unstressed vowels. Neither MSA nor JAA centralizes unstressed vowels, as is the norm in English. These factors work against a tendency of regularity of speech rhythm in MSA and, to a lesser degree, in JAA.

It is highly recommended for future research to continue this endeavor and contrast, on a large scale, rhythm components at the utterance level. Future studies need to pay special attention to the methodology. The accumulated results of such studies would be of great help to Arab learners of English and to English learners of Arabic; this would constitute an important step toward a research-based teaching of pronunciation.

Given the paramount importance of rhythm, the differences between Arabic and English rhythms should be accorded adequate attention. A metalinguistic awareness of these differences along with guided practice would help learners acquire a more stress-timed rhythm. Special focus should be on the most important and most teachable aspects, namely syllable structure, stress, and vowel reduction, together with vowel duration. Arab learners need to be taught to reduce the duration of unstressed vowels and lengthen stressed ones. Conversely, English learners of Arabic need to be trained to slightly shorten stressed vowels and to lengthen unstressed ones.

More attention should be paid to connected speech processes, especially those that are common to both languages such as linking and assimilation. Similarly, Arab learners should be encouraged to use contractions and weak forms in their speech, which will make their rhythm more stress-timed. This would eventually improve their listening skills.

Mohammed Nour Abu Guba (PhD)
University of Sharjah, UAE
mabu-gub@sharjah.ac.ae; abuguba74@gmail.com
ORCID Number: https://orcid.org/0000-0002-5007-6439

Shehdeh Fareh (PhD)
Department of English Language and Literature
University of Sharjah, UAE
shfareh@sharjah.ac.ae
ORCID Number: https://orcid.org/0000-0002-1196-9250

Sane Yagi (PhD)
Department of English Language and Literature

University of Sharjah, UAE
saneyagi@gmail.com
ORCID Number: https://orcid.org/0000-0002-0594-3230
Scopus Code: 24529198100

References

Abercrombie, David. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Abu Guba, Mohammed Nour. (2016). Phonological adaptation of English loanwords in Ammani Arabic. PhD Thesis, University of Salford.

Abu Guba, Mohammed Nour. (2018). 'Stress assignment in polysyllabic words in Levantine Arabic: An optimality-theoretic analysis'. *Lingua Posnaniensis*, 60(2):7–24.

Abu Guba, Mohammed Nour. (2021). 'Gemination within English loanwords in Ammani Arabic: An optimality-theoretic analysis'. *Journal of Linguistics,* 57(1):1–38.

Abu Guba, Mohammed Nour. (Under review). 'Prosody trumps orthography in second language phonology: The case of consonant gemination. *Journal of Psycholinguistic Research*.

Abu Guba, Mohammed Nour. (Under review). 'Foot duration and polysyllabic shortening among Arab speakers of English. *Loquens.*

Abu Guba, Mohammed Nour, Basil Mashaqba, and Anas Huneety. (2023). 'Polysyllabic shortening in Modern Standard Arabic'. *Journal of Semitic Studies*. DOI: fgac030.

Abu Guba, Mohammed Nour, Basil Mashaqba, Samer Jarbou, and Omar Hajeid. (In press). 'Production of vowel reduction among Jordanian Arab learners of English: Acoustic investigation'. *Poznan Journal of Contemporary Linguistics*. **https://doi.org/10.1515/psicl-2022-2011**

Abu Guba, Mohammed Nour, Samer Jarbou, and Abdallah Abu Qub'a. (In press). 'Levels of prosodic prominence in Arabic: The case of function words'. *Al-'Arabiyya.*

Adams, Corinne. (1979). *English Speech Rhythm and the Foreign Learner.* The Hague: Mouton.

Al-Ani, Salman and Darlene R. May. (1973). 'The phonological structure of the syllable in Arabic'. *American Journal of Arabic Studies*, 1:113–125.

Al-Jarrah, Rashid. (2002). An optimality-theoretic analysis of stress in the English of native Arabic speakers. PhD Thesis, Ball State University, USA.

Almbark, Rana, Nadia Bouchhioua, and Sam Hellmuth. (2014). 'Acquiring the phonetics and phonology of English word stress: Comparing learners from different L1 backgrounds'. *Proceedings of the international symposium on the acquisition of second language speech, Concordia Working Papers in Applied Linguistics*, 5:19–35.

Angoujard, J. P. (1990). *Metrical Structure of Arabic*. Dordrecht: Foris publications.

Arvaniti, Amalia. (2009). 'Rhythm, timing and the timing of rhythm'. *Phonetica,* 66(1-2):46–63.

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

Barry, William. (2007). 'Rhythm as an L2 problem: How prosodic is it?'. In Jürgen Trouvain and Ulrike Gut (eds), *Non-Native Prosody: Phonetic Description and Teaching Practice*, 97–120. Berlin: Mouton de Gruyter.

Boersma. Paul, **and David Weenink**. (2015). Praat: Doing phonetics by computer [Computer program]. Version Retrieved July 2019 from http://www.praat.org.

Celce-Murcia, Marianne, Donna Brinton, and Janet Goodwin. (2010). *Teaching Pronunciation.* Cambridge: Cambridge University Press.

Chahal, Dana and Sam Hellmuth. (2014) 'The intonation of Lebanese and Egyptian Arabic'. In Sun Ah Jun (ed.), *Prosodic Typology,* 365–404. Oxford: OUP.

Colantoni, Laura, Jeffrey Steele, and Paola Escudero. (2015). *Second Language Speech: Theory and Practice*. U.K: Cambridge University Press.

Collins, Beverley and Inge Mees. (2008). *Practical Phonetics and Phonology: A Resource Book for Students*. 2nd edition. New York: Routledge.

Cummins, Fred. (2015). 'Rhythm and speech'. In Melissa Redford (ed.), *The Handbook of Speech Production*, 158-177. Malden, Massachusetts: John Wiley and Sons.

Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Segui. (1986). 'The syllable's differing role in the segmentation of French and English'. *Journal of Memory and Language,* 25(4):385–400.

Dalton, Christiane and Barbara Seidlhofer. (1994). *Pronunciation*. Oxford: OUP.

Dauer, Rebecca M. (1983). 'Stress-timing and syllable-timing reanalyzed'. *Journal of Phonetics,* 11:51–62.

Dauer, Rebecca M. (1987). 'Phonetic and phonological components of language rhythm'. *Proceedings of the 11th international congress of phonetic sciences*. Estonia: Academy of Sciences, 447–450.

De Jong, Kenneth and Bushra Zawaydeh. (1999). 'Stress, duration, and intonation in Arabic word-level prosody'. *Journal of Phonetics,* 27(1):3–22.

El-Hassan, Shahir. (1994). 'English accentuation and vowel quality as produced by Arabs: A pedagogic statement'. *Papers and Studies in Contrastive Linguistics,* 29:205–216

Farrah, Mohammed and Nadia Halahlah. (2020). 'Pronunciation problems among Palestinian English major students in Hebron University'. *International Journal of Arabic-English Studies*, 20(1):203–226.

Fletcher, Janet. (2010). 'The prosody of speech: Timing and rhythm'. In William Hardcastle, John Laver, and Fiona Gibbon (eds.) *The Handbook of Phonetic Sciences*, 521–602. Oxford: Wiley-Blackwell.

Ghazali, Salem, Rym Hamdi, and Melissa Barkat. (2002). 'Speech rhythm variation in Arabic dialects'. *Proceedings of the First International Conference on Speech Prosody*, Aix-en-Provence, 331–334.

Ghazali, Salem and Nadia Bouchhioua. (2003). 'The learning of English prosodic structures by speakers of Tunisian Arabic: Word stress and weak forms'. *Proceedings of the 15th ICPhS*, Barcelona, Spain. 961–964.

Ghazali, Salem, Rym Hamdi, and Khouloud Knis. (2007). 'Intonational and rhythmic patterns across the Arabic dialect continuum'. In Elabbas Benmamoun (ed.), *Perspectives on Arabic Linguistics: Papers from the Annual Symposium on Arabic Linguistics,* 97–121.Urbana, Illinois.

Gordon, Matthew. (2011). 'Stress: Phonotactic and phonetic evidence'. In van Oostendorp et al. (eds.), 924–948.

Greenberg, Steven, Hannah Carvey, Leah Hitchcock, and Shuangyu Chang. (2003). 'Temporal properties of spontaneous speech: A syllable-centric perspective'. *Journal of Phonetics,* 31(3-4):465–485.

Gut, Ulrike. (2009). *Non-native Speech: A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Frankfurt: Peter Lang.

Halliday, Michael. (1989). *Spoken and Written Language*. Oxford: OUP.

Hamdi, Rim, Melissa Barkat, Emmanuel Ferragne, and François Pellegrino. (2004). 'Speech timing and rhythmic structure in Arabic dialects: A comparison of two approaches'. *International Speech Conference*, Korea, 1613–1616.

Hayes, Bruce. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago, IL: University of Chicago Press.

Hellmuth, Sam. (2006). Intonational pitch accent distribution in Egyptian Arabic. PhD Thesis. University of London.

Kalaldeh, Raya. (2018). 'Acoustic analysis of Modern Standard Arabic vowels by Jordanian speakers. *International Journal of Arabic-English Studies,*18:23–48.

Kenworthy, Joanne. (1987). *Teaching English Pronunciation*. London: Longman.

Khalifa, Mohammed. (2017). *A Contrastive Metrical Analysis of Main Word Stress in English and Cairene Colloquial Arabic*. Cambridge: Cambridge Scholars Publishing.

Kharma, Nayef and Ali Hajjaj. (1997). *Errors in English among Arabic Speakers: Analysis and Remedy*. Harlow: Longman.

Ladefoged, Peter and Keith Johnson. (2015). *A Course in Phonetics*. 7th edition. Stamford, Connecticut: Cengage Learning.

Lehiste, Ilse. (1972). 'The timing of utterances and linguistic boundaries'. *The Journal of the Acoustical Society of America,* 51(6B):2018–2024.

Miller, M. (1984). 'On the perception of rhythm'. *Journal of Phonetics*, 12:75–83.

Munro, Murray. (1995). 'Nonsegmental factors in foreign accent: Ratings of filtered speech'. *Studies in Second Language Acquisition,* 17(1):17–34.

Munro, Murray and Tracey Derwing. (1995). 'Foreign accent, comprehensibility, and intelligibility in the speech of second language learners'. *Language Learning*, 45(1):73–97.

Odisho, Edward. (2003). *Techniques of Teaching Pronunciation in ESL, Bilingual and Foreign Language Classes*. Munich: Lincom Europa.

Nazzi, Thierry, Josiane Bertoncini, and Jacques Mehler. (1998). 'Language discrimination by newborns: Towards an understanding of the role of rhythm'. *Journal of Experimental Psychology: Human Perception and Performance,* 24:756–766.

Nespor, Marina, Mohinish Shukla, and Jacques Mehler. (2011). 'Stress-timed vs. syllable-timed languages'. In van Oostendorp et al. 1147–1160.

Peperkamp, Sharon and Emmanuel Dupoux. (2002). 'A typological study of stress-deafness'. *Laboratory Phonology*, 7:203–240.

Pike, Kenneth L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan Press.

Prieto, Pilar, and Núria Esteve-Gibert. (2018). *The Development of Prosody in First Language Acquisition*. Amsterdam: John Benjamins Publishing Company.

Rammuny, Raji. (1989). 'Instrumental and auditory analysis of Colloquial Jordanian Arabic: An update'. *Journal of Arabic Linguistics,* 20, 23–42.

Ramus, Franck, Marina Nespor, and Jacques Mehler. (1999). 'Correlates of linguistic rhythm in the speech signal'. *Cognition,* 73:265–292.

Roach, Peter. (1982). 'On the distinction between 'stress-timed' and 'syllable-timed' languages'. In David Crystal (ed.), *Linguistic Controversies. Essays in Linguistic Theory and Practice*. 73–79. London: Arnold.

Roach, Peter. (2000). *English Phonetics and Phonology: A Practical Course*. 3rd edition. Cambridge: Cambridge University Press.

Salem, Nada. (2014). An acoustic investigation of the rhythm of Yemeni Arabic and Jordanian Arabic. PhD Thesis, University of Malaya.

Sluijter, Agaath and Vincent Van Heuven. (1996). 'Spectral balance as an acoustic correlate of linguistic stress'. *The Journal of the Acoustical Society of America,* 100(4):2471–2485.

Soraya, Helmy. (1966). An intonational study of Egyptian colloquial Arabic. PhD Thesis, University of London.

Southwood, Helen and James Flege. (1999). 'Scaling foreign accent: Direct magnitude estimation versus interval scaling'. *Clinical Linguistics and Phonetics*, 13:335–449.

Tajima, Keiichi, Bushra Zawaydeh, and Mafuyu Kitahara. (1999). 'A comparative study of speech rhythm in Arabic, English and Japanese'. *Proceedings of the XIV ICPhS, San Francisco, USA*. 285–288.

Turk, Alice and Stefanie Shattuck-Hufnagel. (2000). 'Word-boundary-related duration patterns in English'. *Journal of Phonetics,* 28(4):397–440.

van Oostendorp, Marc, Colin Ewen, Elizabeth Hume and Keren Rice (eds). (2011). *The Blackwell Companion to Phonology*. Malden, Massachusetts: Wiley-Blackwell.

Watson, Janet C. E. (2011). 'Word stress in Arabic'. In van Oostendorp et al. (eds.), 2990–3018.

Zuraiq, Wael and Joan Sereno. (2007). 'English lexical stress cues in native English and non-native Arabic speakers'. *Ratio*, 1(2):829–832.