

Speech Timing and Rhythmic Structure in Arabic dialects: a comparison of two approaches

Rym Hamdi, Melissa Barkat-Defradas
Emmanuel Ferragne & François Pellegrino

Laboratoire Dynamique Du Langage
UMR CNRS 5596 - Université Lumière Lyon 2 (France)

rim.hamdi@etu.univ-lyon2.fr ; emmanuel.ferragne@univ-lyon2.fr
melissa.barkat@univ-lyon2.fr ; francois.pellegrino@univ-lyon2.fr

Abstract

This paper raises questions about the discrete or continuous nature of rhythm classes. Within this framework, our study investigates speech rhythm in the different Arabic dialects that have been constantly described as stress-timed compared with other languages belonging to different rhythm categories. Preliminary evidence from perceptual experiments revealed that listeners use speech rhythm cues to distinguish speakers of Arabic from North Africa from those of the Middle East. In an attempt to elucidate the reasons for this perceptual discrimination, an acoustic investigation based on duration measurement was carried out (i.e. percentages of vocalic intervals (%V) and the standard deviation of consonantal intervals (ΔC)). This experiment reveals that despite their rhythmic differences, all Arabic dialects still cluster around stress-timed languages exhibiting a different distribution from languages belonging to other rhythm categories such as French and Catalan. Besides, our study suggests that there is no such thing as clear-cut rhythm classes but rather overlapping categories. As a means of comparison, we also used *Pairwise Variability Indices* so as to validate the reliability of our findings.

1. Introduction

It is well known to scholars of Arabic that Arabic vernaculars form a bipolar continuum with the Middle East representing the Eastern Dialects and North Africa hosting the Western Dialects. Although this division is an oversimplification of Arabic dialectology, it is widely accepted by the linguistic community and maybe supported by native speakers' behavior. In the present paper, we have verified the notion of the continuum. All Arabic dialects, generally classified as having stress-timed rhythm, exhibit rhythmic structures that vary geographically from North Africa to the Middle East. Based on the computation of percentages of vocalic intervals (%V) and the standard deviation of consonantal intervals (ΔC), we showed in a previous study that these differences seemed to be connected with variations of the rhythmic structure [1]. In this paper we compare the rhythmic variation in these different dialects usually classified as stress-timed with the rhythm of other languages belonging to other categories. The purpose of this comparison is to find out whether the stress timing of Arabic dialects is a discrete category based on syllable types or if rhythm should be regarded as a continuum.

2. Discrete Rhythmic Classes and the notion of Rhythmic Continuum

In speech, rhythm has been defined as an effect involving the isochronous recurrence of some type of speech unit [2] [3] [4]. All these authors suggested that all spoken languages are either stress-timed (e.g. English, Arabic...) or syllable-timed (e.g. French, Spanish...) or mora-timed (e.g. Japanese, Tamil...). According to [5] measurements obtained from the acoustic signal could not provide evidence for classifying languages (including Arabic) as stress-timed, as opposed to syllable-timed, if stress-timing meant isochrony of inter-stress intervals. Dauer [6] supported this view showing that interstress intervals were no more regular in English, a stress-timed language, than in Spanish, a syllable-timed language. She came to similar conclusions and proposed a new system for rhythmic classification claiming that all languages are more or less stress-based. She argued that the perception of different types of rhythm has mainly to do with differences in syllable structure, vowel reduction and types of stress. On the strength of this statement, [7] proposed that perceived rhythm classes corresponding to the phonological properties put forth in [6] and [8] could be accounted for by instrumental measurements of the acoustic signal, which we will explain later with respect to Arabic speech rhythm in the following section.

3. Rhythm in Arabic

In the literature on Arabic speech rhythm, all the dialects investigated have been consistently categorized as 'stress-timed' as opposed to 'syllable-timed' or 'mora-timed' languages [3] [9] [10] [11]. Evidence for these classifications comes essentially from perceptual experiments. Moreover, [13] reported that, during a listening task, her subjects – who were linguistically naïve native speakers of Arabic from various regions in the Arab world – were able to identify successfully Arabic speakers as belonging to North Africa or the Middle East 98% of the time. On the same task, native speakers of French with no knowledge of Arabic were able to discriminate between the two regions 56% of the time only. These results were reported to be statistically significant. When asked to list the criteria that had helped them make their decisions, most subjects mentioned that North African Arabic sounded faster, and jerkier than Eastern Arabic an impression that, if translated into phonetical terms would be correlated with speech rhythm. Now, if different languages may be perceived as belonging to

different rhythm classes, and if Arabic is always categorized as a stress-timed language, what cues in the acoustic signal made the listeners felt that Western Arabic sounded more *staccato* than Arabic in the Middle East? In other words, can a language or a dialect be more or less stress-timed than another ?

4. Method

In this study we investigate the production of rhythm in six Arabic dialects and three other languages. Arabic dialects were chosen from different geographic areas: Moroccan, Algerian and Tunisian representing Western Arabic (WA), Jordanian Syrian and Egyptian representing the Eastern Arabic (EA). Besides, we added three other languages belonging to other traditional rhythm classes: English as stress-timed, French as syllable-timed and Catalan as an intermediate language [7]. The speech data were taken from free translations of the story: "*the North Wind and the Sun*". Each Arabic subject listened sentence by sentence to the story presented in French before translating spontaneously each sentence into his native dialect. Catalan and American subjects translated the same data from French into their mother tongue while French subjects translated it from English. We recorded three male speakers from each language and/or dialect. The language corpus used in this paper consists in 270 sentences (i.e. 10 sentences \times 9 languages \times 3 subjects per language) with an average duration of 2.5 seconds for each sentence. We used a signal processing and editing program to identify and classify segments as vowels or consonants. Our first method consisted in measuring syllabic complexity and diversity following the experimental procedures proposed by [7]. We computed (i) the duration of each sentences, (ii) the duration of each string of consecutive vowels (vocalic intervals), (iii) the duration of each string of consecutive consonants (consonantal intervals). The next step consisted in computing the proportion of vocalic and consonantal intervals (V% and C% respectively) in each sentence, and the standard deviation of vocalic and consonantal intervals within each sentences (ΔV and ΔC respectively). Based on their findings [7] argued that a combination of %V and ΔC provided the best acoustic correlates of rhythm classes. In stress-timed language, which has full and reduced vowels, %V was smaller than in syllable-timed language, which does not have vowel reduction. Moreover, ΔC was larger in stress-timed language and reflect the more complex syllable. This is the reason why we chose to work with the combination of these two parameters only.

Though this approach seems to be reliable for the discrimination of languages and/or dialects where syllabic complexity varies significantly. We thought it would be interesting to compare the results obtained with the combination of V% ΔC , a method corresponding to a global representation of speech rhythm according to [12], with a procedure based on a more local approach of speech rhythm, that is the *Pairwise Variability Indices* (henceforth *PVI*) [14]. This technique – which measures the mean difference in duration between two successive vowels over a whole utterance – is expected to be more sensitive to the stress pattern exhibited in the language. Though the V% ΔC and the *PVI* methods both aim at describing rhythm characteristics

across languages, we wanted to compare them in order to see whether the parameters involved are correlated.

5. Results

5.1. The V% ΔC method

Cross-dialectal comparison shows that while the proportion of vocalic intervals represents less than 50% of the total duration of a sentence in all Arabic dialects, it is more important in the Eastern dialects than in their Western counterparts. In fact, there is a gradual increase of %V as one moves from West to East. Conversely, ΔC decreases from West to East. Figure 1 illustrates the negative correlation between %V and ΔC and dialect location ($r = -0.75$). As one moves from West to East ΔC decreases and %V increases. Table 1 and Figure 1 below show the results for all languages and dialects in our corpus. French subjects present higher values of vocalic intervals than other languages and Arabic dialects. The standard deviation of consonantal intervals in French seems to be similar to those observed in Eastern dialects.

Table 1: Computed values for %V and ΔC

Area	Language/Dialect	%V	ΔC
Western	Moroccan	30.81	102.22
	Algerian	31.14	57.54
	Tunisian	34.86	57.54
Eastern	Egyptian	36.82	52.38
	Jordanian	40.48	55.06
	Syrian	39.76	51.10
Other languages	French	43.38	48.72
	English	38.07	75.27
	Catalan	38.00	60.37

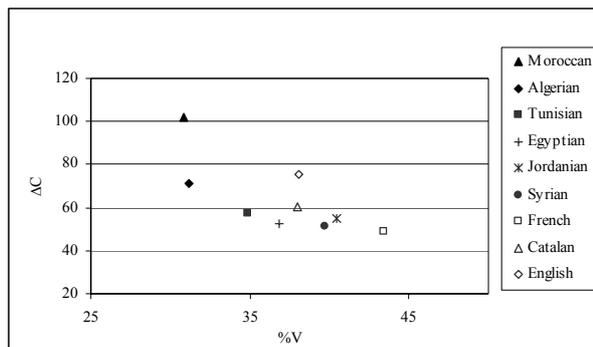


Figure 1: Distribution of languages & dialects along the %V (x axis) and ΔC dimensions (y axis)

Using t-test for the difference between pairs of dialects, we found that the significance level is directly proportional to the distance between dialects and languages. For example, while these results (for both %V and ΔC) are not significant when Syria is compared to Jordan or when we compared Morocco to Algeria that is, when two dialects belong to the same region, they are highly significant for pairs of dialects located in the opposite ends of the continuum such as Syrian and Moroccan ($p < 0.001$). Figure 2 is an illustration of the average values

when the three dialects of each region are grouped together. It clearly shows that %V is higher in the Eastern Arabic dialects than in Western Arabic dialects ($p < 0.0001$), while the opposite results is obtained for ΔC . Results show the data on western dialects that have closer interval to English: language that have been cited as prototypical example of stress, than French: syllable-timed language.

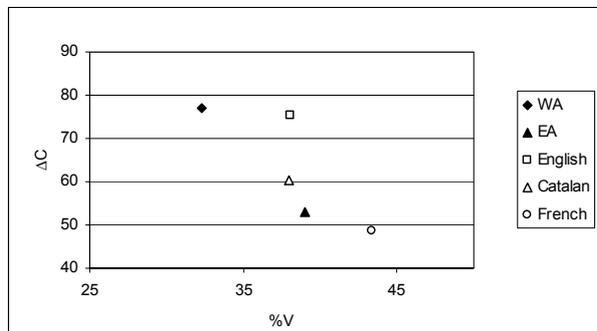


Figure 2: Comparison of %V (x axis) and ΔC (y axis) in Western Arabic WA , Eastern Arabic EA and in the other languages.

5.1. The PVI method

As for the correlation between the values obtained with %V \sim ΔC and PVI methods we observe that both ΔC and RPVIC (corresponding to the raw consonant-duration pairwise variability index), and ΔV and RPVIV (corresponding to the raw vowel-duration pairwise variability index) are highly correlated ($r^2 = 0.83$ and $r^2 = 0.72$ respectively) (Figures 3 & 4). This suggests on the one hand that using either RPVIC or ΔC basically comes down to the same thing. On the other hand, given that RPVIV was found to be a fairly reliable cue for the discrimination of languages [14] and that it is strongly correlated with ΔV , we may infer that the latter may be useful for our purpose although [7] concluded that it was not relevant for his data. The next step for us will be to compute RPVIV and RPVIC so as to compare the results with those obtained using %V \sim ΔC combination.

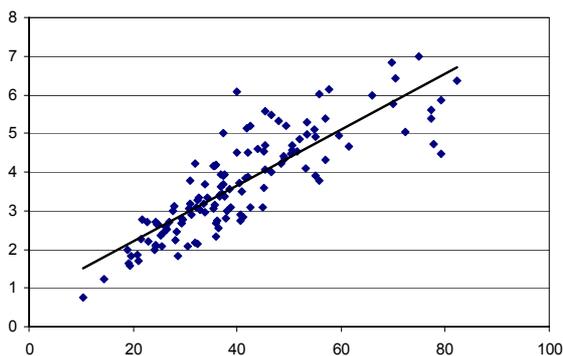


Figure 3 : Scatter plot & linear regression for RPVIV (x axis) ΔV (y axis)

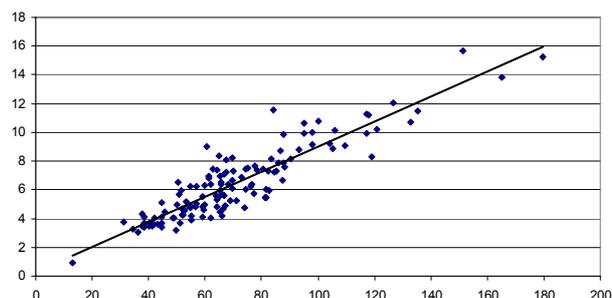


Figure 4 : Scatterplot & linear regression for RPVIC (x axis) ΔC (y axis)

6. Discussion & Conclusion

Lower proportions of vocalic intervals reflect the presence of shorter vowels. All investigations of Arabic vowel space have underlined the fact that in North African dialects, phonologically long and short vowels have shorter duration than the corresponding vowels in the dialects of the Middle East. The standard deviation of consonantal intervals is correlated with syllabic diversity and complexity as revealed by [7]. North African Arabic is well known for processes of short vowel deletion in open syllable resulting in various consonants clusters and types of syllables with complex onsets and codas. The complex syllables coupled with reduced vowels, especially in the Moroccan dialects resulting in the impression of the jerky and halting speech reported in [13]. Differences in vowel duration and syllabic complexity seem to be the main factors responsible for differences in rhythmic structures. Languages with the highest ΔC and the lowest %V such as English were those traditionally classified as stress-timed. In our study, the dialects that exhibit these characteristics are those spoken in North Africa. Since Eastern dialects such as those of Iraq and Jordan have also been classified as stress-timed [11]. On the basis of these results, Arabic dialects can be more or less stress-timed or syllable-timed, we should perhaps expect a great deal of variation within the class of stress-timing. Therefore, a strict categorical distinction between stress-timing and syllable-timing cannot be defended. Indeed, according to [7], and in order to maintain a discrete 'stress-timed' category as distinct from some other timing, there should exist one or more key factors the presence of which regularly induces the perception of stress-timing. Such a conditioning factor could be the tendency in all Arabic dialects for long or heavy syllables to attract stress [8]. Note also how dialects geographically located between the two poles are also intermediate with respect to phonetic reality (Figure 1). Barkat [15] reported that most of the discrimination errors made by her subjects concerned Tunisian speakers. In fact, Tunisian dialect attest a %V similar to North Africa but a ΔC closer to the Middle East. In other words, their vowels are slightly longer and less reduced than those of Moroccans and Algerians, but significantly shorter than Syrian and Jordanian. They don't, however, exhibit the same syllabic complexity as the other North African subjects. The comparison between Arabic values and French confirms the notion of a continuum for rhythmic categories instead of absolute classes. In fact, the proportion of vocalic values in all Arabic dialects is more reduced than those observed in French. Besides, ΔC French values occur to be comparable to

those observed in Eastern dialects, that is the latter exhibit a syllabic structure comparable to that of French (i.e. predominance of open syllables) as opposed to the complex syllabic structures attested in Western Arabic dialects characterized by typical consonantal clusters. These results confirm the existence of a large rhythmic class where a great amount of variation was found even if the comparison is carried out between languages traditionally belonging to different classes. In order to develop this notion, a useful perceptual task may be to ask non native speakers of Arabic to classify synthesized speech data both from different Arabic dialects and other languages known to have different types of stress-timing and see whether the same pattern of discrimination emerge. This part of work is under process.

7. References

- [1] Ghazali, S., Hamdi, R. & Melissa Barkat-Defradas, Speech Rhythm variation in Arabic dialects, in proceedings of the 1st international Conference on Speech Prosody, Aix-en-Provence, 2002, 331-334.
- [2] Abercrombie, D. Elements of General Phonetics, Edinburgh University Press, 1967.
- [3] Pike, K; *The Intonation of American English*. 2nd edition. Ann Arbor: University of Michigan Press. 1945.
- [4] Bloch, B. Studies in colloquial Japanese IV: Phonemics. *Language* 26, 1950, pp. 86-125.
- [5] Roach, P. On the distinction between 'stress-timed' and 'syllable-timed' languages. In *Linguistic controversies*, Crystal Ed, London: Edward Arnold, pp. 73-79. 1982.
- [6] Dauer, R.M. Stress-timing and syllable-timing reanalysed, *Journal of Phonetics*, 1983, pp 51-69
- [7] Ramus, F., Nespor, M. & Melher, J. Language identification with suprasegmental cues. A study based on speech resynthesis. *Journal of the acoustical Society of America* 105(1), 1999, 512-521.
- [8] Dauer, R.M., Phonetic and phonological components of language rhythm, *Proceedings of the XIth ICPhS*, Tallinn, Estonia, 1987, vol.5, pp. 447-450.
- [9] Miller, M. "On the perception of rhythm", *Journal of Phonetics*, 12, 1984, pp. 75-83.
- [10] Benguerel, A. Stress-timing vs syllable-timing vs mora-timing: The perception of speech rhythm by native speakers of different languages. *VARIA, Etudes & Travaux*, vol 3, 1999.
- [11] Tajima, K. Zawaydeh, B. Kitabara, M. A comparative study of speech rhythm in Arabic , English and Japanese, *Proceedings of the XIV ICPhS*, San Francisco, U.S.A, 1999.
- [12] Ferragne, E., Pellegrino, F., Rhythm in Read British English: interspeaker and interdialect variability, submitted to *Interspeech* 04.
- [13] Barkat, M., Ohala, J-J. , Pellegrino, F., Prosody as a distinctive feature for the discrimination of Arabic dialects, in *Proceedings of Eurospeech* Budapest, 1999, pp. 395-398.
- [14] Grabe, E. and Low, E.L. Durational Variability in Speech and the Rhythm Class Hypothesis. *Papers in Laboratory Phonology* 7, Mouton, 2002.
- [15] Barkat-Defradas M, Vers l'identification automatique des parlers arabes, *revue Langues & Linguistique*, vol. 7, pp. 47-95