Rhythm in L2 speech

Rytm w mowie L2

Ulrike Gut

University of Münster, Germany
gut@uni-muenster.de

ABSTRACT

This chapter explores to what extent the current conceptualizations and measurements of speech rhythm can be employed in research on second language speech. It discusses why many of the quantitative rhythm metrics, due to their sensitivity to speech rate, cannot be used to measure the rhythm of L2 speech reliably. Furthermore, it investigates whether L2 speech rhythm can be meaningfully compared to rhythm in native speech and whether stages of acquisition of speech rhythm can be described with these quantitative rhythm metrics. Finally, the question is raised which validity these measurements of L2 speech rhythm have and which phonological properties of speech are actually measured by them.

STRESZCZENIE

W niniejszym rozdziale analizujemy do jakiego stopnia istniejące obecnie ujęcia i pomiary rytmu mowy mogą być zastosowane w badaniu mowy w drugim języku. Omawiamy dlaczego wiele spośród ilościowych metrum rytmu, z uwagi na czułość względem tempa mowy, nie może być zastosowanych do rzetelnego pomiaru rytmu w L2. Ponadto, badamy czy rytm mowy w L2 może być w sposób znaczący porównany z rytmem w mowie natywnej i czy stadia nabywania rytmu mowy mogą być opisane za pomocą tych ilościowych metrum rytmu. Na zakończenie pytamy o to, jaką zasadność mają te pomiary rytmu w mowie w L2 i które fonologiczne własności mowy są przez nie mierzone.

1. Speech rhythm and rhythm metrics

The term 'speech rhythm' refers to the temporal organization of language. Following Lloyd [1], Pike [2] and Abercrombie [3], speech rhythm has been conceptualised as a periodic recurrence of events and as a structural property of languages by which they can be allocated to different rhythm classes or be located along a rhythmic continuum. Abercrombie [3:p96] for example, proposed that in syllable-timed languages such as French syllables "recur at equal intervals of time – they are isochronous" [3:p97]. In stress-timed languages such as English, by contrast, the foot, the time interval between two stress beats is assumed to be isochronous, i.e. roughly equal in time.

Although it was shown to be impossible to confirm these early claims of isochrony empirically [4–9], a battery of different rhythm metrics have been developed in order to capture systematic quantitative rhythmic differences between languages. Ramus et al. [10], for example, proposed to divide speech into vocalic and consonantal parts and to compute the proportion of the vocalic intervals of a sentence (*%V*), the standard deviation of vocalic intervals (*delta V*) and the standard deviation of consonantal intervals (*delta C*). Comparing five read sentences in eight languages along the axes of the measured percentage of vowels and the standard deviation of the consonantal intervals, they showed that English, Dutch and Polish cluster together with a high *delta C* and relatively low *%V*. French, Spanish, Catalan and Italian group together with a lower *delta C* and a higher *%V*, while Japanese has the lowest *delta C* and the highest *%V*. Since it was found that *delta C* is inversely related to speech rate [11, 12], Dellwo [13] introduced the rate-normalised metric *VarcoC*, which computes the standard deviation of consonantal interval duration divided by the mean consonantal duration and multiplied by 100. By the same token, *VarcoV* calculates the standard deviation of vocalic interval duration divided by the mean vocalic interval duration and multiplied by 100.

Other rhythm metrics do not measure global properties of utterances, but rather focus on the degree of durational contrast between neighbouring events. Low & Grabe [14] and Grabe and Low [15] calculate speech rhythm with the *Pairwise Variability Index* (*PVI*), which computes the sum of the durational differences between adjacent vocalic or consonantal intervals in an intonation phrase, taking the absolute value of each difference and dividing it by the duration of the pair. When analysing passages read by one speaker each in different languages, Grabe & Low found that Thai, Dutch, German, British English, Tamil and Malay group together with regard to their high variation in vowel durations, whereas Mandarin and Spanish show a low variation in vowel length. The *consonantal PVI* is lowest for Estonian, Romanian, French, Welsh and Luxemburgish and highest for Polish, Catalan, Tamil, Malay and British English. In order to control for speech rate variation, Low, Grabe & Nolan [16] introduced the *normalised Pairwise Variability Index* (*nPVI*), which calculates the mean absolute normalised difference between durations of neighbouring interval pairs; normalisation was performed by dividing each duration difference by the mean of the durations concerned (1).

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

In yet another approach to measuring rhythm, Deterding [17] describes differences between British and Singapore English rhythm based on the *Variability Index* (*VI*), which calculates the mean durational differences between subsequent syllables. Alternatively, Gibbon & Gut [18] propose the *Rhythm Ratio* (*RR*), which calculates the average ratio of all adjacent syllable or vowel pairs as a percentage (multiplied by 100); cf. example (2).

(2) $$RR = 100 \times \left[\sum_{k=1}^{m-1} rr_k \right] / (m-1)$$

$(\textit{if } d_k < d_{k+1} \textit{ then } rr_k = d_k/d_{k+1} \textit{ else } rr_k = d_{k+1}/d_k)$

Gut, Adouakou, Urua and Gibbon [19] compared the speech rhythm of the three West African languages Anyi, Ega and Ibibio with the *Rhythm Ratio* and found that the West African languages group closely together and show conspicuous differences in both successive syllable durations and successive vowel durations compared to British English. A variant of the *Rhythm Ratio* is the *Syllable Ratio*, in which only the average ratio of adjacent syllable pairs of the type stressed:unstressed (full-vowelled:reduced-vowelled) is calculated [20].

Although cross-linguistic comparisons of speech rhythm based on these metrics show that languages can be localised as closer towards one or the other end of a continuum between syllable-timing and stress-timing, it has been repeatedly questioned whether these metrics with their focus on vowel, consonant and syllable duration actually succeed in measuring speech rhythm. Cummins [21], for example, proposes that rhythm is located on a phonological rather than the phonetic level and that it consists of a coupling between intervals at one prosodic level to those at a higher one. Several experiments showed that speakers of different languages are varyingly successful in repeating a phrase aligned with a rhythmic acoustic signal: While speakers of English aligned the stressed syllables of phrases with the beats given by a metronome, speakers of Spanish and Italian did not [22]. Furthermore, timing patterns differed between English and Arabic speakers on the one hand and Japanese speakers on the other [23].

Other critics have proposed that the concept of speech rhythm is merely used as a cover term for a range of structural properties of a language. Dauer, for instance, suggested that "rhythmic differences […] between languages […] are more a result of phonological, phonetic, lexical, and syntactic facts about that language than any attempt on the part of the speaker to equalize interstress or intersyllable intervals" [9:p55]. She points out that speech rhythm measured in syllable duration differences merely reflects syllable structure distribution, which varies enormously across languages. Equally, differences in rhythm between languages reflect whether a language has vowel reduction or not; those classified as stress-timed usually do. In addition, syllable-timed languages either do not have lexical stress, or accent is realised by variations in pitch contour. Conversely, stress-timed languages realise word level stress by a combination of length, pitch, loudness and quality changes, which result in clearly discernible beats. Consequently, Barry [24], argues for second language teaching to dispense with the concept of speech rhythm altogether and to instead concentrate on practising the underlying phonological processes such as vowel reduction in weak forms, vowel-length and vowel-quality contrasts and consonant-cluster reductions.

Given these reservations concerning quantitative phonetic measurements of speech rhythm and given their increasing popularity in the investigation of non-native (L2) speech, this contribution sets out to discuss some of the fundamental questions concerning L2 speech rhythm. In detail, these are:

1. Can L2 speech rhythm be measured reliably with quantitative rhythm metrics? Especially since many of them appear to be sensitive to speech rate, the slower speech rate of non-native speech might distort the measurements.
2. Can L2 speech rhythm be meaningfully compared to rhythm in native speech?
3. Can the acquisition of speech rhythm be described with these quantitative rhythm metrics?
4. Is the measurement of L2 speech rhythm valid? If yes, which phonological correlates of speech rhythm are being measured?

These questions will be addressed in the following section. Section 3 gives a conclusion and outlook.

2. Rhythm in L2 speech

The past few years have seen an enormous increase in the number of investigations concerned with the rhythmic properties of L2 speech (e.g. [25, 20, 26–36], and a wide array of different rhythm metrics have been employed in these studies. Yet, how reliable are these instruments? Are measurements of the rhythmic properties of languages stable and can they be replicated across different studies?

2.1. Reliability of rhythm metrics

It appears that most rhythm metrics do not yield similar results across different studies. The *%V* values calculated for British English, for example, are 45.7% in Arvaniti [37], 41.1% in Grabe & Low's study [15], 40.1% in Ramus et al.'s study [19] and 38% in White and Mattys analysis [33]. The *delta C* value for British English is given in Grabe & Low as 56.7 and as 53.5 in Ramus et al., but it was found to be 60 in Arvaniti's study. Moreover, the *delta C* value for German varies from 52.6 [15] to 71.7 [12]. The vocalic nPVI values for Thai given in Grabe & Low [15] and Sarmah et al. [36] are 65.8 and 54.5 respectively. The vocalic *nPVI* for Spanish is given as 20 by Thomas & Carter [32], while it is given as 36 by White & Mattys [33]. In addition, the vocalic *nPVI* measured for British English is given as 57.2 in Grabe & Low and 73 in White & Mattys. Grabe & Low's value is in fact very close to White & Mattys's value for French (50).

This observed unreliability of the rhythm metrics is probably largely due to the different segmentation procedures used in different studies as well as the influence of speaking style and material selection. All of the rhythm metrics calculate fine-grained durational differences between vowels, consonants or syllables, but the segmentation procedure of the acoustic signal varies considerably across studies: for example, formant transitions for obstruents were counted as part of the vowel in Thomas & Carter [32:p340], but as part of the consonantal interval by Gut [28:p70]. Likewise, post-vocalic glides were counted as part of the vowel in Ramus et al. [10] but as consonants by Arvaniti [37] if they showed evidence of frication. Likewise, syllabic consonants were counted as vowels in Thomas & Carter [32:p342] but not in other studies such as Gut [28:p70].

As already predicted by Jassem [38:p429], it has furthermore been shown that the rhythm values of languages vary considerably across different speaking styles. In a large-scale study, Arvaniti [37] demonstrated for six different languages that all rhythm metrics give significantly different values when participants are reading individual sentences as opposed to reading a story. Equally, Thomas & Carter [32] measured a vocalic *nPVI* of 20 for spontaneous Spanish, while it is given as 36 in reading passage style in [33]. Likewise, [39] found that syllable durations varied significantly less in spontaneous speech than in reading passage style in Swedish – a difference, however, that was not found for Spanish.

Furthermore, the selection of test sentences can influence the resulting rhythm values: *nPVI* scores derived from alternating patterns and monotonic geometric series are the same if they have the same ratio, so that, for example, sentences with the durational pattern (2,4,2,4,2,4) and the pattern (2,4,16,32,64) have equal *nPVI* values (in these cases 100), as Gibbon [40] has pointed out. This was proven empirically by Arvaniti [37], who measured in six languages the rhythm of especially designed 'stress-timed' sentences, which incorporated as much variability as the language allowed, and 'syllable-timed' sentences, which included simple syllable structures and little consonantal or vocalic variability. She found that in all languages rhythmic differences between the two sets of sentences were significant and concludes that rhythmic variability within a language can be as high as rhythmic variability across languages. Moreover, in addition to the type of selected sentences, their number may influence resulting rhythm values. Studies range in terms of the amount of material used from five sentences per speaker [10, 33] to more than 80.000 syllables in the corpus-based analysis of L2 rhythm [28].

While these factors influencing the reliability of rhythm metrics might be controlled for by employing consistent segmentation procedures and selecting comparable material (but see Arvaniti [37], who doubts this), it is the relationship of these measurements with speech rate that is most compromising for the reliability of analysing rhythm in L2 speech. A number of the rhythm metrics are significantly correlated with speech rate and thus yield different measurements depending on the speaker's speed of articulation. The metrics *delta C* and *delta V*, for example, have been repeatedly shown to be inversely related to speech rate [11, 33]: the faster the speech, the lower the measured variability will be. Likewise, Deterding [17] observed that the correlation between speech rate and the VI was negative in both Singapore English and British English. The faster his participants spoke the smaller durational differences were found between adjacent syllables. The *Syllable Ratio* is also significantly correlated with articulation rate in both English and German [28:p192]. Confusing results were found for the *vocalic PVI*. Thomas & Carter [32:p347], report no effect of speech rate for African American English and American English, but found one for Spanish and Jamaican English: the *vocalic PVI* increased with increasing syllable duration.

It thus appears that these rhythm metrics should only be employed in comparing L2 speakers with each other or with L1 speakers when the speech rate of all speakers is controlled for. Some metrics, however, seem to be relatively independent of a speaker's speech rate: Little correlation between *%V* and speech rate has been found so far [12, 33:p509], which makes this the most reliable rhythm metric for L2 speech. Similarly, the normalised variance measures *VarcoV* and *VarcoC* compensate for speech rate variations. Yet, do these metrics succeed in differentiating between native and L2 speech rhythm?

2.2. Differentiation of L1 and L2 speech

One fundamental question to be addressed when analysing L2 speech rhythm is whether it can be meaningfully differentiated from L1 rhythm by using the proposed rhythm metrics. Can these measurements of rhythm reliably describe differences between native and non-native speakers of a language? It appears that this is only partly the case. When a speaker's first and second language are rhythmically similar, none of the metrics seem to be able to differentiate between native and non-native speech rhythm: White & Mattys, for instance, compared the rhythm of English learners of Dutch and Dutch learners of English – all of them with “a discernible non-native accent“ [33:p506] – with respective native speakers but found no significant differences measured with any of the metrics *delta V*, *delta C*, *%V*, *VarcoV*, *VarcoC*, *vocalic nPVI* and *consonantal PVI*.

It might of course be argued that in speakers acquiring a second language that is rhythmically similar to their first language no difference between native and non-native speech rhythm is to be expected because learners can employ positive transfer from their L1's rhythm patterns. But even for speakers acquiring a second language with distinctly different rhythmic properties from their first language, the analysis of rhythmic properties is far from straightforward: The *vocalic PVI* metric, for instance, gave rhythm values for Spanish L2 speakers of English that were significantly different from both L1 Spanish and L1 American English in Thomas & Carter's study [32:p345], but White & Mattys [33] found that both *consonantal PVI* and *vocalic PVI* cannot differentiate between either native and non-native English or native and non-native Spanish rhythm. Stockmal et al. [31] report that these two metrics cannot distinguish between non-native and native speakers of Latvian either. Yet, they differ significantly between Korean learners of English and native speakers [35]. Likewise, while some Korean learners of English show a higher %V than native speakers [35], this metric does not yield significantly different values for Japanese learners of English [34]. *Delta V*, similarly, gives fairly variable results: While the rhythm of Spanish spoken as an L1 and an L2 (by English native speakers) differs significantly in this measurement, the same metric cannot differentiate between English spoken by native speakers and spoken by Spanish L1 speakers [33].

Conversely, some metrics have been shown to reflect differences between L1 and L2 speech rhythm: Both Spanish and Korean learners of English show a lower vocalic variability – measured with *VarcoV* – than native speakers [33, 35]. Similarly, native Canadian English speakers of French have a significantly higher *Variability Index* in spontaneous speech than native speakers of European French [25]. The *Syllable Ratio*, furthermore, succeeded in displaying significant differences between native and non-native German as well as native and non-native English speech rhythm [28].

Other metrics that have at least in some studies been found to differentiate between native and non-native speech are those measuring consonantal variability, *delta C* and *consonantal PVI*. Nigerian speakers of English have higher *delta C* values than British English speakers [26]. Japanese speakers of Canadian English have a higher consonantal variability – measured as *consonantal PVI* – than Canadian English native speakers [34], just as low-proficiency native Russian speakers of Latvian compared to Latvian native speakers do [31]. By the same token, Chinese, Italian and English L2 speakers of German display increased *delta C* values in comparison to native Germans [28], and *delta C* differs significantly between non-native Russian and native speakers of Latvian [31]. Both *delta C* and the *consonantal PVI* are also higher in English produced by Korean L1 speakers than in native English speech [35]. Lin and Wang [30], conversely, found that advanced Chinese speakers of L2 English have *delta C* scores that are not different from those of native English speakers. Neither did White & Mattys [33] or Jang [35] find significant differences in *delta C* between native and non-native rhythm. Since *delta C* has been shown to correlate with speech rate, higher *delta* C values in L2 speech rhythm have to be interpreted with extreme caution: It is quite possible that they simply reflect the lower speech rate of individual second language speakers. This view is corroborated by Payne, Post, Astruc, Prieto & del Mar Vanrell's findings [41], who showed that the speech of two-, four- and six-year-olds has significantly higher *delta C* and *consonantal PVI* values than adult speech, too.

One factor that might be responsible for some of these contradictory findings in the investigation of systematic differences between L1 and L2 speech rhythm is that of the L2 speakers' level of competence. The speech of near-native learners of a language is to be expected to show different rhythmic properties than the speech of relative beginners and consequently the selection of language learners in studies that compare L2 rhythm with L1 rhythm will influence the results. Yet, are the rhythm metrics actually able to differentiate between beginners' and more advanced language learners' speech rhythm? Can they trace rhythmic developments in language acquisition?

2.3. Measuring the acquisition of L2 rhythm

Not many studies so far have been concerned with the development of rhythmic properties in L2 speech, and very few of them so far have put forward evidence for rhythmic differences related to a language learner's level of competence. The rhythm of beginners, intermediate and advanced Korean learners of English showed no significant difference when measured with either *%V*, *delta V*, *delta C*, *vocalic nPVI*, *consonantal PVI*, *VarcoV* or *VarcoC* [35]. Guilbault [25:p119] compared rhythmic properties of spontaneous speech of more and less experienced learners of French with Canadian English as their L1. While the less experienced learners' *Variability Index* was lower than that of the more experienced learners, the difference between the two groups was not significant. Neither the L2 proficiency nor the time spent in the target language environment seems to have had any significant influence on their speech rhythm. The same was found by Gut [28], who compared the rhythm – measured with the *Syllable Ratio* – of language learners before and after a 9-month stay abroad in the target language country or a 6-month training course in pronunciation. No effect of either the training or the stay abroad was found in terms of rhythmic properties. Similarly, when Sarmah et al. [36] compared the rhythm of Thai speakers of English who had spent fewer than six months in the US at the time of recording with Thai speakers who had been living there for more than 18 months, they found that the speakers with less exposure showed a higher *nPVI* and a lower *%V* values than the learners with more exposure, who thus exhibited values that were further distant from native English measures.

Yet, in one study at least some rhythm metrics appear to be able to differentiate between learners at different competence levels. Significant differences in *delta C* and *consonantal PVI* are reported between beginners and more advanced learners of Latvian with Russian as their L1 [31]. *%V* and *vocalic PVI* values, by contrast, did not distinguish between the two groups of learners.

One fundamental question arising from these inconsistent findings is that of the validity of the rhythm metrics. If it appears impossible to measure differences in speech rhythm between language learners at different competence levels and if rhythmic developments cannot be traced with them, what is it then they actually measure? Do these metrics actually succeed in differentiating different types of speech rhythm? Or to put it in other words: Do these measurements really capture rhythm or do they rather measure other properties of speech?

2.4. Validity of rhythm measurements

The first study to test the validity of a wide range of rhythm metrics were White & Mattys [33], who compared five sentences read by six speakers each of Spanish, French, English and Dutch. The seven metrics they employed were *delta C*, *delta V*, *%V*, *VarcoV*, *VarcoC*, *vocalic nPVI* and *consonantal PVI*. Of these, only *%V*, *vocalic nPVI* and *VarcoV* succeeded in tracing both the rhythmic differences between more stress-timed Dutch and English on the one hand and more syllable-timed French and Spanish on the other and the gradient differences within these groups. Arvaniti [37], however, who analysed the rhythm of English, German, Greek, Italian, Korean and Spanish by applying these rhythm metrics to read sentences – some of which were manipulated to be more stress-timed or more syllable-timed as described in section 2.1 above –, a read story and spontaneous speech of eight speakers each, concluded that these "metrics cannot be reliably used to classify languages into rhythm classes" (p. 28 of pre-print article). This view is based on the observed effects that different speaking styles, test materials and speakers have on these metrics. The type of elicitation method, the composition of test materials and inter-speaker variation caused a variability in durational timing within each of the languages that was larger than the variability across them. Consistent cross-linguistic rhythm differences, according to Arvaniti, therefore cannot be captured with these quantitative duration-based metrics.

Is it then the case that these rhythm metrics do not actually measure rhythm? The analysis of the findings on L2 speech rhythm indeed seems to confirm Dauer's [9] and Barry's [24] view that the concept of 'speech rhythm' may just be a label for a combination of phonological processes such as vowel reduction and consonant cluster reduction that apply in a language. This is for example suggested by the studies that found that L2 speech, just like child speech [41] has a higher variability in consonantal intervals than native (adult) speech [31, 34, 28] and that this variability decreases with increasing level of competence [31]. What more does this reflect than the learners' initial difficulty with consonant (cluster) articulation and their subsequent increasing mastery of coarticulatory and reduction processes during language acquisition? A mastery that results in a faster overall articulation rate, a process that might be described as 'becoming more fluent'? In fact, Jang [35], who measured the speech of native English speakers and Korean learners of English at three competence levels with seven rhythm metrics, found that none of them was able to distinguish between native and non-native speech on the one hand and the different learner groups on the other as well as the simple measurement of speech rate [35:p50]. It seems quite probable then that *delta C* and the *consonantal PVI* – which are known to correlate with speech rate – are just an alternative way of measuring a learner's speed of articulation.

Quite a number of 'unexpected' rhythm values can furthermore be explained by vocalic processes in L2 speech. In Sarmah et al.'s study [36], for example, L2 English produced by Thai native speakers unexpectedly displayed a higher vocalic *nPVI* than British English native speech (59.6 vs. 57.2) despite Thai having a lower one (54). This 'overshoot', the fact that there is even greater durational variability between vowels in Thai L2 English than in native English, seems to be a result of very specific differences in intrinsic vowel length between Thai English and native English: Sarmah et al. found that the mean length of the vowels /ɪ/ and /ʊ/ are considerably shorter in Thai English (on average 106 ms) than in American English (on average 192 ms).

By the same token, White & Mattys [33] observed an unexpected greater *%V* for English L2 speakers of Spanish than for native Spanish speakers. Their explanation of this finding is that the English natives incorrectly produced diphthongs instead of monophthongs in some places such as the last vowel in Spanish ‘coche’. In addition, the authors hypothesised that the English speakers might have produced a pitch accent on this word, which would have led to a lengthening of the vowel. Gut’s study [28] serves as a third example: She found that English learners of German had a lower *%V* value than native speakers and Italian and Chinese learners of German. A detailed analyses of vowels in the pretonic unstressed syllable ver- and the posttonic unstressed syllables of the type C+<-en>, C+<-e> and C+<-em> showed that the English learners, unlike the native speakers of German, produced reduced vowels in the pretonic unstressed syllables; in addition, they significantly more frequently deleted vowels in the posttonic unstressed syllables. The measured differences in ‘rhythm’ thus can be explained and described by systematic vocalic and intonational processes in L2 speech.

3. Conclusion and outlook

This chapter has shown the difficulty of measuring the rhythm of L2 speech. Most of the metrics that are in use for comparing native speech rhythm across languages can neither distinguish between native and non-native speech rhythm – particularly if the learner’s native and target language are rhythmically similar – nor yield significantly different values for L2 learners at different competence levels. All of them are moreover influenced by the choice of speakers, materials and type of speaking style, as Arvaniti [37] has shown. Most crucially, when describing language learners’ rhythm and when comparing it to native speech rhythm with the help of these metrics, it has to be borne in mind that the resulting values reflect not only methodological choices but also the differences in speech rate between speakers. This makes a comparison of learner groups at different competence levels (which are nearly always characterised by differences in speaking rate) and the investigation of rhythmic development practically impossible and a comparison with native speakers very difficult. The resulting values of the latter comparison could only be interpreted meaningfully if the native speakers were asked to produce speech at the same rate as the non-native speakers.

What is more, phonological characteristics of L2 speech that have been interpreted as rhythmic properties can be easily explained and described with reference to other phonological processes such as consonant cluster reduction and vowel reduction processes. It appears that what has been termed ‘rhythm’ and what is measured by the quantitative rhythm metrics in the speech of L2 learners are just the phonetic by-products of these phonological processes as Barry [24] has suggested. One conclusion to be drawn from this is that these metrics do not really capture the essence of rhythm [18] and that other approaches are necessary for the investigation of L2 speech rhythm. These could be approaches that conceptualise rhythm on a phonological rather than a phonetic level, such as the ones proposed by Jassem [38] or Cummins [21]. Maybe what language learners have to acquire is an abstract conception of a language-specific timing pattern which is then flexibly applied to the varying segmental composition of different utterances. Some support for this comes from a study reported by Missaglia [42]: She developed the *Contrastive Prosody Method* for Italian learners of German, in which the learners first practise the delivery of L1 utterances with the L2 prosody, i.e. the production of Italian utterances with only one strong accent, thus deaccenting and reducing all the other words. Learners who had mastered the production of Italian utterances with a German ‘rhythm’ were then able to transfer these prosodic patterns to German utterances.

One other conclusion to be drawn from the failure of the rhythm metrics to reveal the properties of L2 rhythm is of course that it does not exist in the first place. Given the evidence from the studies carried out so far, it is probably time to ask whether it is really useful to assume that L2 speech (and child speech for that matter) has any kind of systematic rhythm. If what has traditionally been referred to as speech rhythm is simply the sum of the durational features of consonantal and vocalic processes influenced by speech rate, what is the added scientific benefit of proposing the domain of rhythm? Thus, the research findings on L2 rhythm might tempt one to go even a step beyond Arvaniti [37]. While she argues that there is little validity in the construct of rhythm classes, one might now – roughly 60 years after its ‘invention’ by Lloyd [1], Pike [2] and Abercrombie [3] – want to raise the fundamental question of abandoning the concept of speech rhythm altogether.

References

[1] Lloyd James, A. 1940. *Speech Signals in Telephony*. London: Sir I. Pitman & Sons.

[2] Pike, K. 1945. *The Intonation of American English*. Ann Arbor: University of Michigan Press.

[3] Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

[4] O’Connor, J. D. 1965. The perception of time intervals. Phonetics Laboratory, University College London, Progress Report 2, 11–15.

[5] Uldall, E. 1971. Isochronous stresses in R.P. In L. Hammerich, R. Jacobson and E. Zwirner, eds., *Form and Substance*. Copenhagen: Akademisk Forlag. 205–210.

[6] Hill, D., W. Jassem, and I. Witten. 1979. A statistical approach to the problem of isochrony in spoken British English. In H. Hollien and P. Hollien, eds., *Current Issues in Linguistic Theory* 9: *Current Issues in the Phonetic Sciences*. Amsterdam: John Benjamins. 285–294.

[7] Fauré, G., D. Hirst, and M. Chafcouloff. 1980. Rhythm in English: isochronism, pitch, and perceived stress. In L. Waugh and C. van Schooneveld, eds., *The melody of Language*. Baltimore: University Park Press. 71–79.

[8] Roach, P. 1982. On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages. In D. Crystal, ed., *Linguistic Controversies, Essays in Linguistic Theory and Practice*. London: Edward Arnold. 73–79.

[9] Dauer, R. 1983. Stress-timing and syllable-timing reanalysed. *Journal of Phonetics* 11, 51–62.

[10] Ramus, M. Nespor and J. Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292.

[11] Barry, W. J., B. Andreeva, M. Russo, S. Dimitrova and T. Kostadinova, T. 2003. Do rhythm measures tell us anything about language type? In *Proceedings of the 15th international congress of phonetics sciences*, Barcelona. 2693–2696.

[12] Dellwo, V. and P. Wagner. 2003. Relations between language rhythm and speech rate. In *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona. 471–474.

[13] Dellwo, V. 2006. Rhythm and speech rate: A variation coefficient for deltaC. In P. Karnowski, & I. Szigeti, eds., *Language and language processing: Proceedings of the 38th linguistic Colloquium, Piliscsaba* 2003. Frankfurt: Peter Lang. 231–241.

[14] Low, E. L. and E. Grabe. 1995. Prosodic patterns in Singapore English. *Proceedings of 13th International Congress of Phonetic Sciences, Stockholm, Sweden*. 636–639.

[15] Grabe, E. and E. L. Low. 2002. Durational variability in speech and the rhythm class hypothesis. In: C. Gussenhoven and N. Warner, eds., *Papers in Laboratory Phonology* 7. Berlin: Mouton. 515–546.

[16] Low, E. L., Grabe, E. & Nolan, F. (2000). Quantitative characterisations of speech rhythm: 'Syllable-timing' in Singapore English. *Language and Speech*, 43, 377–401.

[17] Deterding, D. 2001. The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics* 29, 217–230.

[18] Gibbon, D. and U. Gut. 2001. Measuring Speech Rhythm. In *Proceedings of Eurospeech, Aalborg*. 91–94.

[19] Gut, U., S. Adouakou, E.-A. Urua and D. Gibbon. 2002. Rhythm in West African tone languages: a study of Ibibio, Anyi and Ega. In U. Gut and D. Gibbon, eds., *Typology of African Prosodic Systems, Bielefeld Occasional Papers In Typology (BOPIT)* 1. Bielefeld: Universität Bielefeld. 159–165.

[20] Gut, U. 2003a Non-native speech rhythm in German. *Proceedings of 15th International Congress of Phonetic Sciences, Barcelona*. 2437–2440.

[21] Cummins, F. 2002 Speech Rhythm and Rhythmic Taxonomy. In Bernard Bel and Isabelle Marlien, eds. *Proceedings of Speech Prosody 2002*. Aix-en-Provence, Laboratoire Parole et Langage. 121–126.

[22] Cummins, F. and R. Port. 1998 Rhythmic constraints on stress-timing in English. *Journal of Phonetics* 26, 145–171.

[23] Tajima, Keiichi, B. Zawaydeh, M. Kitahara and R. Port. 2000. Finding acoustic evidence for speech rhythm across languages: A speech cycling approach. *Journal of the Acoustical Society of Japan* 21, 287–289.

[24] Barry, W. J. 2007. Rhythm as an L2 problem. How prosodic is it? In J. Trouvain and U. Gut, eds. *Non-native Prosody: Bridging the Gap between Research and Teaching*. 97–120.

[25] Guilbault, C. 2002. *The Acquisition of French Rhythm by Second Language Learners*. PhD thesis, University of Alberta.

[26] Gut, U. 2003b. Nigerian English – a typical West African language? In Ewald Mengel, H.-J. Schmid and M. Steppat, eds., *Proceedings of the Anglistentag 2002 Bayreuth*. Trier: Wissenschaftlicher Verlag Trier. 461–471.

[27] Gut, U. 2005. Nigerian English prosody. *English World-Wide* 26, 153–177.

[28] Gut, U. 2009. *Non-native Speech: a Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Frankfurt: Peter Lang.

[29] Carter, P. M. 2005. Quantifying Rhythmic Differences between Spanish, English, and Hispanic English. In R. Gees, ed., *Theoretical and Experimental Approaches to Romance Linguistics: Selected Papers from the 34th Linguistic Symposium on Romance Languages*. Amsterdam: John Benjamins. 63–75.

[30] Lin, H. and Q. Wang. 2005. Vowel quantity and consonant variance: A comparison between Chinese and English. In *Proceedings of Between Stress and Tone. Leiden, June 2005*.

[31] Stockmal, V., D. Markus, and D. Bond. 2005. Measures of Native and Non-Native Rhythm in a Quantity Language. *Language and Speech* 48, 55–63.

[32] Thomas, E. R. and P. M. Carter. 2006. Prosodic rhythm and African American English. In *English World-Wide* 27, 331–355.

[33] White, L. and S. L. Mattys. 2007. Calibrating rhythm: First language and second language studies. In *Journal of Phonetics* 35, 501–522.

[34] Grenon, I. and L. White. 2008. Acquiring rhythm: A comparison of L1 and L2 speakers of Canadian English and Japanese. In H. Chan, H. Jacob and E. Kapia. *Proceedings of the 32nd Annual Boston University Conference on Language Development*. Somerville: Cascadilla. 155–166.

[35] Jang, T.-Y. 2008. Speech Rhythm Metrics for Automatic Scoring of English Speech by Korean EFL Learners. In *Malsori Speech Sounds The Korean Society of Phonetic Sciences and Speech Technology* 66, 41–59.

[36] Sarmah, P., D. V. Gogoi and C. Wiltshire. 2009. Thai English. Rhythm and vowels. In *English World-Wide* 30, 196–217.

[37] Arvaniti, A. In press. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*.

[38] Jassem, W. 1983. *The Phonology of Modern English*. Warsaw: Państwowe Wydawnictwo Naukowe.

[39] Engstrand, O. and D. Krull. 2003. Rhythmic intentions or rhythmic consequences? Cross-language observations of casual speech. *Proceedings of the 15th International Congress of the Phonetic Sciences, Barcelona*. 2789–2792.

[40] Gibbon, D. 2003. Computational modelling of rhythm as alternation, iteration and hierarchy. In *Proceedings of the 15th international congress of phonetic sciences, Barcelona*. 2489–2492.

[41] Payne, E., B. Post, L. Astruc, P. Prieto and M. del Mar Vanrell. 2011. Measuring Child Rhythm. *Language and Speech* 54, 1–27.

[42] Missaglia, F. 2007. Prosodic training for adult Italian learners of German: the Contrastive Prosody Method. In J. Trouvain and U. Gut, eds., *Non-native prosody. Phonetic description and teaching practice*. Berlin: Mouton de Gruyter. 237–258.