Durational Variability in Speech and the Rhythm Class Hypothesis

Esther Grabe*

Department *of Linguistics, University of Cambridge, Cambridge CB3 9DA, U.K.*

Address for correspondence: Phonetics Laboratory, 41 Wellington Square, Oxford OX1 2JF

Ee Ling Low

National *Institute of Education, Nanyang Technological University, English Language and Literature, 1 Nanyang Walk, Singapore 637616*

Abstract

In this paper, we provide evidence for rhythmic classifications of speech from duration measurements. Our investigation differs from previous studies in two ways. Firstly, we do not relate speech rhythm to phonological units such as interstress intervals or syllable durations. Instead, we calculate durational variability in successive acoustic-phonetic intervals using *Pairwise Variability Indices*. Secondly, we compare measurements from languages traditionally classified as stress-, syllable- or mora-timed with measurements from hitherto unclassified languages. The values obtained agree with the classification of English, Dutch and German as stress-timed and French and Spanish as syllable-timed: durational variability is greater in stress-timed languages than in syllable-timed languages. Values from Japanese, a mora-timed language, are similar to those from syllable-timed languages. But previously unclassified languages do not fit into any of the three classes. Instead, their values overlap with the margins of the stress-timed and the syllable-timed group.

1. Introduction and Background

In the present paper, we investigate the acoustic-phonetic basis of speech rhythm. In speech production, rhythm has been defined as an effect involving the isochronous recurrence of some type of speech unit, a view made popular by Pike (1946) and Abercrombie (1965, 1967). Pike and Abercrombie suggested that all spoken languages exhibit isochronous units of speech, and that languages are either stress-timed or syllable-timed. In stress-timed languages, intervals between stresses or rhythmic feet are said to be near-equal, whereas in syllable-timed languages, successive syllables are said to be of near-equal length. In Pike and Abercrombie's view, the distinction between stress- and syllable-timing was strictly categorical; languages could not be more or less stress- or syllable-timed. Abercrombie (1965) based his categorical distinction on the physiology of speech production. All spoken languages were said to have two types of pulses, chest-pulses and stress-pulses. Chest pulses were pulse-like puffs of air from the lungs, resulting from alternate contractions and relaxations of the breathing muscles. Stress-pulses were less frequent, more powerful contractions of the breathing muscles which reinforce some of the chest-pulses. Rhythm, Abercrombie suggested, was a product of the way in which the two pulse-systems combined. Two categorically different combinations were possible (Abercrombie, 1965). In syllable-timing, chest-pulses were in isochronous sequence, but stress-pulses were not. In stress-timing, stress-pulses re-enforced chest pulses in isochronous sequence.

A third type of rhythm, mora-timing, was proposed by Bloch (1942), Han (1962), and Ladefoged (1975). Mora-timing was exemplified by Japanese. Traditionally, morae are sub-units of syllables consisting of one short vowel and any preceding onset consonants. In mora-timing, successive morae are said to be near-equal in duration. Thus, mora-timed languages are more similar to syllable-timed languages than to stress-timed languages.

1.1 *Evidence for stress- and syllable-timing from duration measurements*

The empirical basis of the rhythm class hypothesis has been investigated extensively, but experimental support for isochrony in speech is lacking (Beckman, 1992, Laver, 1994). In stress-timed languages, interstress intervals are far from equal, and interstress-intervals do not pattern more regularly in stress-timed than in syllable-timed languages (Shen and Peterson, 1962, Bolinger, 1965, Delattre, 1966, Faure, Hirst and Chafcouloff 1980, Pointon, 1980, Wenk and Wioland, 1982, Roach 1982, Dauer, 1983, Manrique and Signorini, 1983, Nakatani, O'Connor and Aston, 1981, Dauer, 1987, Eriksson, 1991). Nor are syllables or morae of roughly equal length in syllable-timed languages (Pointon, 1980, Wenk and Wioland, 1982, Roach 1982, Dauer, 1983, 1987). Roach (1982), for instance, compared interstress intervals in languages classified as stress-timed and languages taken to be syllable-timed. He investigated two claims made by Abercrombie (1967) about the difference between stress-timed and syllable-timed rhythm: (i) there is considerable variation in syllable length in a language spoken with stress-timed rhythm, whereas in a language spoken with syllable-timed rhythm, syllables tend to be equal in length, and (ii) in syllable-timed languages, interstress intervals are unevenly spaced. Roach's findings did not support either claim. The syllable-timed languages in his sample exhibited greater variability in syllable durations than the stress-timed languages. Roach also observed a wider range of percent deviations in interstress intervals in stress-timed than in syllable-timed languages. Roach concluded that measurements of time intervals in speech could not provide evidence for rhythm classes. Roach's view has been supported by Dauer's (1983) study. Dauer compared interstress intervals in English, Thai, Spanish, Italian and Greek. She found that interstress intervals were no more regular in English, a stress-timed language, than in Spanish, a syllable-timed language. Dauer concluded that the search for acoustic phonetic correlates of stress- and syllable-timing was futile.

Isochrony in mora-timing was investigated by Han (1962) Port, Al-Ani and Maeda (1980), and Port, Dalby and O'Dell (1987). Port et al. (1987) argue that these studies provide some preliminary support for the mora as a constant time unit. But other researchers have questioned the acoustic basis for mora-timing (Oyakawa, 1971, Beckman, 1982, Hoequist, 1983a,b). Beckman (1982)'s data, for instance, did not show that segments vary in length in Japanese in order to compensate for intrinsic durations of adjacent segments so that morae are equal in length.

In short, although popular among linguists, the rhythm class hypothesis has been contradicted by numerous empirical studies. Abercrombie's view of speech rhythm as a combination of chest and stress-pulses has long been disproven (e.g. Ladefoged, 1967), destroying the physiological basis of a strict categorical distinction into stress- and syllable-timed languages. The predictions for speech timing arising from the rhythm class hypothesis have suffered a similar fate. Researchers have not provided support from duration measurements for isochronous timing, on any absolute basis (Laver, 1994). This failure has obliged some researchers to retreat from 'objective isochrony' to 'subjective isochrony'. These researchers describe the physical regularity of isochrony as a tendency (Beckman, 1992, Laver, 1994). True isochrony is assumed to be an underlying constraint, and the surface realisation of isochronous units are perturbed by phonetic, phonological and grammatical characteristics of the language. Other researchers have concluded that isochrony is primarily a perceptual phenomenon (e.g. Lehiste 1977, Couper-Kuhlen 1990, 1993). Proponents of the 'isochrony-as-perception' view argue that the differences in duration measured between interstress-intervals or syllable durations are well below the threshold of perception. Consequently, isochrony may be accepted as a concept that relates to speech perception.

1.2 *Other views of rhythm in speech*

The weak empirical evidence for isochrony led Dauer (1983, 1987) to propose a new system for rhythmic classification. In Dauer's view, speakers do not attempt to equalise interstress or intersyllable intervals. Instead, all languages are more or less stress-based. Dauer suggests that prominent syllables recur at regular intervals in English, a stress-timed language, but also in Spanish, a syllable-timed language. But in English, prominent syllables are perceptually more salient than in Spanish. Consequently, rhythmic diversity results from the combinations of phonological, phonetic, lexical and syntactic facts associated with different languages. Syllable-structure, the presence or absence of vowel reduction, and word stress are especially relevant to rhythmic differences. In stress-timed languages, syllable structures are more varied than in syllable-timed languages. In syllable-timed languages, vowel reduction is rarely found. Dauer's account is similar to a proposal published a year earlier by Dasher and Bolinger (1982). Dasher and Bolinger suggested that the rhythm of a language is the result of specific phonological phenomena such as variety of syllable types, the presence or absence of phonological vowel length distinctions, and vowel reduction.[1] Dasher and Bolinger argued that rhythm type is not a phonological primitive but results from the phonological structure of a given language.

Nespor (1990) offered another view of speech rhythm that differs from the continuous system proposed by Dauer. Nespor argued against traditional rhythmic categories on the basis of rhythmically intermediate languages. Intermediate languages exhibit some properties associated with stress-timing and some associated with syllable-timing. Nespor held that neither a dichotomous view nor a continuous classification system can adequately account for the rhythmic properties of such languages. The languages she cited to support her proposal were Polish, which has been classified as stress-timed but does not exhibit vowel reduction, and Catalan, which has been described as syllable-timed but has vowel reduction.

1.3 *Evidence for rhythm classes from durational variability*

The present study concerns the relationship between speech timing and rhythmic classifications of languages. We depart, however, from the search for isochrony.[2] We did not measure interstress intervals or syllable durations which are phonological units. Some of their phonetic correlates can be disputed. Instead, we followed Low, Grabe and Nolan (2000) and took a direct route from impressionistic observations of rhythmic differences between languages to the acoustic signal. We measured the durations of vowels, and the duration of intervals between vowels (excluding pauses) in a passage of speech. Then we computed a Pairwise Variability Index for each type of measurement. The index expresses the level of variability in successive measurements. The raw Pairwise Variability Index (*rPVI*) is given in equation (1).

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

where m is number of interrvals, vocalic or intervocalic, in the text and d is the duration of the kth interval. Notice that rPVI is not normalised for speech rate.

Low et al. used a normalised version of the Pairwise Variability Index in their measurements on vowel durations. The equation for this version, the normalised Pairwise Variability Index (*nPVI*), is

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

where *m* is number of items in an utterance and *d* is the duration of the *k*th item.

Equation (2) shows that the *nPVI* is compiled by calculating the difference in duration between each pair of successive measurements, taking the absolute value of the difference and dividing it by the mean duration of the pair. Equation (1) for the *rPVI* differs only in omitting the third step. The differences are then summed and divided by the number of differences. The output is multiplied by 100, because the normalisation produces fractional values.

In previous studies (Low and Grabe, 1995, Low, 1998, Grabe, Post and Watson, 1999, Low, Grabe and Nolan, 2000), we applied the *nPVI* to vowel durations.[3] This followed work by authors such as Beckman and Edwards (1994) who showed that vowels constitute the lowest level of the prosodic hierarchy (at least in English). Our studies revealed that so-called stress- and syllable-timed languages differ in the durational variability encountered in vowels. Stress-timed languages such as English exhibit more vocalic variability than syllable-timed languages such as French (Grabe, Post and Watson, 1999). We related this finding to vowel quality. English has full as well as spectrally reduced and shortened vowels. The consequence is a high level of variability in vowel durations. French does not have vowel reduction, and the level of vocalic variability is significantly lower. Low et. al (2000) applied the *nPVI* to data from ten speakers of British English (stress-timed) and ten speakers of Singapore English (syllable-timed). The data provided an acoustic basis for the impression of syllable-timing in Singapore English. Statistical analyses showed that vowel durations are significantly more variable in British English than in Singapore English. Deterding (1994) obtained similar results in an investigation of spontaneous speech data from British English and Singapore English.

In a related study, Ramus, Nespor and Mehler (1999) set out to provide acoustic evidence for the traditional stress-timing/syllable-timing dichotomy. Ramus and colleagues argued that a viable account of speech rhythm should not rely on complex and language-dependent phonological concepts but on purely phonetic characteristics of the speech signal. These authors segmented speech into vocalic and consonantal intervals. In other words, they measured vowel durations and the duration of intervals between vowels. Ramus et al. computed three acoustic correlates of rhythm from the measurements: (a)*%V*, the proportion of time devoted to vocalic intervals in the sentence, disregarding word boundaries; (b) ΔV : the standard deviation of vocalic intervals; (c) ΔC : the standard deviation of consonantal intervals, the sections between vowel offset and vowel onset.

On the basis of their findings, Ramus et al. argued that a combination of *%V* and ΔC provided the best acoustic correlate of rhythm classes. In English, which has full and reduced vowels, *%V* was smaller than in French, which does not have vowel reduction. On the other hand, ΔC was larger in English and reflected the more complex syllable options available in that language.

Low et al. (2000) compared the *nPVI* with the standard deviation measures ΔV and ΔC. The authors concluded that a Pairwise Variability Index may be a better indicator of rhythmicity than ΔV or ΔC. Ramus and colleagues measured duration in a set of tightly controlled sentences from eight languages (5 sentences each produced by four speakers). In less tightly controlled data, Low and colleagues argued, the standard deviation would reflect spurious variability introduced by changes in speaking rate within and across sentences and between-speaker differences in speaking rate. Consider a language where three successive long vowels follow three successive short vowels and another where long and short vowels alternate. Both would give the same standard deviation, although the pattern of vowel durations differs radically between the two.

Low et al. (2000) concluded their paper by suggesting an addition to the vocalic *nPVI*. The standard deviations published by Ramus et al. (1999) showed that rhythmically mixed languages such as Catalan and Polish exhibit complementary levels of vocalic and intervocalic variability. In Polish, the standard deviation of vocalic intervals was relatively low, making Polish similar to the syllable-timed languages in the sample. But the standard deviation of intervocalic intervals was comparatively high. The reverse applied to Catalan. Low and colleagues suggested that a combination of their vocalic *nPVI* with a measure of intervocalic interval variability would provide a better indicator of rhythmic class than the vocalic *nPVI* alone. This combination would capture the rhythmic characteristics of stress-timed, syllable-timed and mixed languages. Low and colleagues predicted that English (stress-timed) should exhibit relatively high variability index values for vocalic and intervocalic intervals. Some English syllables are relatively complex and we find consonant clusters in the onset and in the coda. Others have a very simple structure. Consequently, intervocalic variability is likely to be high. Spanish (syllable-timed) should have low values in both types of interval. Successive vowels are similar in length, and a large proportion of syllables have a simple CV structure (Dauer, 1983). Polish (mixed) would be low on the vocalic axis and high on the intervocalic axis. Catalan (mixed) would be high on the intervocalic axis, and low on the vocalic axis.

In the present paper, we have tested the predictions made by Low, Grabe and Nolan. We also have added measurements on a range of rhythmically unclassified languages. Our aim was to establish whether the unclassified languages would pattern with the stress-timed or the syllable-timed group or whether some or all of them would be intermediate.

2. Method

2.1 *Languages*

We made duration measurements on comparable passages of speech from eighteen languages (one speaker per language). The subjects read the ‘North Wind and the Sun’, a standard text from phonetic research. Translations of this text into Catalan, Dutch, English, French, German, Japanese, and Thai are available in the Handbook of the International Phonetic Association (1999), accompanied by brief phonetic and phonological analyses. Translations not available in the handbook were made by the subjects and by colleagues. In Table 1, we list the languages investigated and, where possible, their rhythmic classification.

Language	**Classification**
British English	Stress-timed (Classe, 1939, Pike, 1946, Abercrombie, 1967)
German	Stress-timed (Kohler, 1982)
Dutch	Stress-timed (Ladefoged, 1975, Smith, 1976)
Thai	Stress-timed (Luangthongkum, 1977)
Tamil	Syllable-timed (Corder, 1973, Asher, 1985)
Spanish	Syllable-timed (Pike, 1946, Hockett, 1958)
French	Syllable-timed (Abercrombie, 1967, Catford, 1977)
Singapore English	Syllable-timed (Tongue, 1979, Platt and Weber, 1980)
Japanese	Mora-timed (Bloch, 1942, Han, 1962)
Polish	Mixed (Dauer, 1987, Nespor, 1990)
Catalan	Mixed (Dauer, 1983, Nespor, 1990)
Estonian	Unclassified
Greek	Unclassified
Luxembourg	Unclassified
Malay	Unclassified
Mandarin	Unclassified
Rumanian	Unclassified
Welsh	Unclassified

Table 1: Traditional rhythmic classifications of languages investigated in the present study.

2.2 *Recording procedure*

We recorded one speaker from each language. The British English, French, German, Greek, Polish, Rumanian, and Welsh subjects were recorded in a sound-treated booth in the Oxford Phonetics Laboratory. Dutch, Estonian, Japanese, Luxembourg, and Thai subjects were recorded in a comparable booth in the Cambridge Phonetics Laboratory. The Catalan data were recorded in a sound-treated room at University College London. Tamil, Malay, and Singapore English were recorded in a quiet room at the National Institute for Education in Singapore. The Mandarin data were also recorded in Singapore and represent the variety of Mandarin spoken in Singapore. The Spanish data were provided by Anders Eriksson (University of Stockholm). Subjects were asked to read the text once, at their own pace. They were given time to read the text before the recordings were made.

2.3 *Acoustic analysis*

Duration measurements were made by the first author, with assistance from a colleague. Vocalic intervals were defined as the stretch of signal between vowel onset and vowel offset, characterised by vowel formants, regardless of the number of vowels included in the section (a vocalic section could contain a monophthong, a diphthong, or, in some cases, two or more vowels spanning the offset of one word and the onset of the next). Intervocalic intervals were defined as the stretch of signal between vowel offset and vowel onset, regardless of the number of consonants included. The duration of vocalic and intervocalic intervals was measured left-to-right, using wide-band spectrograms generated by xwaves™. Vowels were identified using generally accepted criteria (Peterson and Lehiste, 1960, Fischer-Jørgensen and Hutters, 1981). For instance, in fricative-vowel sequences, the onset of the vowel was taken to be the onset of the second formant. In vowel-voiceless fricative sequences, the vowel was considered terminated where the noise pattern began. In vowel-voiced fricative sequences, we considered the vowel terminated at the onset of high frequency energy. Nasal-vowel sequences were segmented by observing the fault transitions between nasal and vowel. Our approach to glides was based on acoustic, not phonetic or phonological criteria. In initial glides, the formant movements continue seamlessly from glide to vowel. We excluded initial glides from vocalic portions if their presence was indicated by clearly observable changes in formant structure or in the amplitude of the signal. Otherwise, glides were included in the vocalic portion.

Pauses between intonation phrases, as well as hesitations, were excluded from the analysis. One consequence of this approach is that some intervals taken to be continuous are, in fact, split by a pause. Although this approach is not ideal, it allows us to calculate *PVI* values from longer samples of speech. In earlier work, we calculated *PVI* values for individual intonation phrases (Low and Grabe, 1995, Grabe et al., 1999, Low et al., 2000). In the present paper, we have departed from our earlier approach for three reasons. Firstly, we wished to distance ourselves from as many subjective or intuitive linguistic decisions as possible when taking measurements. The location of intonation phrase boundaries is debatable, and native speakers of a language can disagree on where a boundary should be placed. In the present study, we investigated some languages that we do not speak. Consequently, we could not determine the location of intonation phrase boundaries in these languages with any certainty. Secondly, in earlier work on British English and Singapore English, we excluded intonation phrase-final syllables from the index, because we knew that Singapore English is characterised by relatively more phrase-final lengthening than British English. In the present study, we have no information on the relative degree of phrase-final lengthening or shortening in the languages investigated, and therefore we include phrase-final syllables. Finally, the number of intervals separated by a pause is relatively small and is not likely to change the results greatly.[4]

Finally, a note on vowel devoicing: English and Japanese, for instance, exhibit devoiced vowels. In spectrograms, devoiced vowels do not exhibit the formant patterns which characterise voiced vowels. Consequently, as our approach to measuring is acoustic, not phonological, we have measured the duration of a vowel only if there was evidence of a voiced vowel in the acoustic signal.

2.4 *Normalisation*

In previous work, we applied the *nPVI* to vocalic intervals, normalising for changes in speaking rate. In the present paper, we have investigated intervocalic as well as vocalic intervals, so we have reconsidered the question of normalisation. A significant correlation between interval duration and speaking rate would support the application of normalisation. For our purposes, we defined speaking rate as the average vocalic or intervocalic interval duration produced by a speaker. We examined the effect of speaking rate on interval variability across speakers using data from (a) the twenty speakers who provided data for Low et al., and (b) data from the present study (one speaker of each of eighteen languages). The effect of speaking rate on the *rPVI* (*PVI* before normalisation) was tested across speakers and across languages. In British English, vocalic and intervocalic *rPVI* values increased significantly as the average interval duration increased across speakers (Pearson's r, vocalic intervals: r=0.867, p<0.01, intervocalic intervals: r=0.801, p<0.01, 2-tailed). The data from the present study (one speaker per language) showed similar results (vocalic intervals: r=0.613, p<0.01, intervocalic intervals: r=0.808, p<0.01, 2-tailed).

These results confirm that normalisation is desirable for vocalic intervals. The results for intervocalic intervals suggested that we may also need to normalise for intervocalic interval duration. However, this is not a necessary conclusion. Vocalic and intervocalic intervals differ in one respect, which affects the rate normalisation question. An intervocalic interval is compositional. It can contain several different segmental units, and these may be subject to different speech rate effects. But the majority of our vocalic intervals consists of a single vowel that is stretched or compressed when speech rate changes. If we consider that intervocalic units can hold roughly one to five segments in Polish but only one to two segments in Japanese, we can account for part of the correlation between intervocalic intervals and speech rate across languages. The correlation arises from (a) an increase in the intervocalic *rPVI* as the duration of intervocalic intervals increases, and (b) cross-language differences in syllable structure. Languages that exhibit a greater number of syllable-structure options may be associated with a greater intervocalic *rPVI*. The correlation between the duration of intervocalic intervals and language may be due to the combined effects of speaking rate and cross-language differences in syllable-structure. These effects are not easy to tease apart. An investigation into the details of rate normalisation in intervocalic intervals goes beyond the scope of the present paper. In what follows, we continue to normalise vocalic intervals for speech rate, as in previous work.[5] For intervocalic intervals, we do not apply the normalisation, using the *rPVI*.[6]

2.5 Predictions

We predicted that stress-timed languages would exhibit high vocalic *nPVI* and high intervocalic *rPVI* values. Syllable-timed languages would have low vocalic *nPVI* and low intervocalic *rPVI* values. Polish, a mixed language with complex syllable structure and no vowel reduction was predicted to exhibit a lower vocalic *nPVI* value than stress-timed languages but a relatively high intervocalic *rPVI* value. Catalan, another mixed language, was expected to have a relatively high vocalic *nPVI* value combined with a low intervocalic *rPVI*, possibly similar to the intervocalic *rPVI* of Spanish. We did not make any predictions for Estonian, Greek, Luxembourg, Malay, Mandarin, Rumanian and Welsh. Japanese (mora-timed) has a relatively simple syllable structure. As there is no vowel reduction in Japanese, we predicted that the vocalic *nPVI* would be similar to the vocalic *nPVI* values of syllable-timed languages (e.g. French or Spanish). The relatively simple syllable-structure led us to expect a low intervocalic *rPVI* also. However, between voiceless consonants, vowels are often devoiced and not associated with formant patterns. In our measurements, devoiced vowels were included in intervocalic intervals, and intervocalic intervals containing a devoiced vowel were longer than intervocalic intervals separated by voiced vowels. This approach necessarily raises the intervocalic *rPVI* value for Japanese.

3. Results

3.1 *PVI Results*

Figure 1 shows the data on languages that have often been cited as prototypical examples of stress-, syllable- and mora-timing: British English, Dutch and German (stress-timed), French and Spanish (syllable-timed) and Japanese (mora-timed). Vocalic *nPVI* values are plotted on the vertical axis against intervocalic *rPVI* values on the horizontal axis.

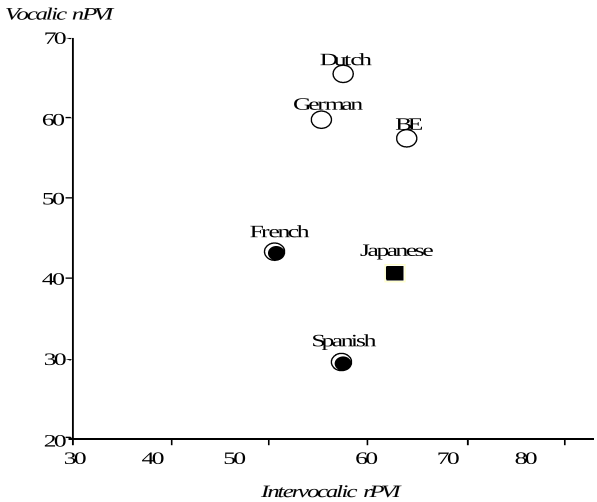

Figure 1. *PVI* profiles from prototypical stress-timed languages English, Dutch and German, syllable-timed languages French and Spanish, and mora-timed language Japanese. Vocalic variability is plotted on the vertical axis against intervocalic variability on the horizontal axis. ○= stress-timed, ●= syllable-timed, ■ = mora-timed.

The *PVI* profiles provide acoustic evidence for rhythmic differences between English, Dutch and German on the one hand, and French and Spanish on the other. English, Dutch and German have been described as stress-timed and exhibit high vocalic *nPVI* values. French and Spanish have been described as syllable-timed and exhibit low vocalic *nPVI* values. This finding supports the rhythmic classification suggested by Pike (1946) and Abercrombie (1967), even if the evidence does not come from isochronous interstress-intervals or syllable-durations. There is no support, however, for a strict categorical distinction between languages with high vocalic and intervocalic *PVI* values and languages with low vocalic and intervocalic *PVI* values. Rather, it appears that languages can be more or less ‘stress-timed’ or ‘syllable-timed’.

Figure 1 also shows that our predictions for the intervocalic *rPVI* are supported by the contrast between French (syllable-timed), and the British English, Dutch and German (stress-timed). French, which has a relatively simple syllable structure, appears to have a lower intervocalic *rPVI* than English, Dutch and German, which have more complex syllable-structures. Spanish, however, exhibits a lower intervocalic *rPVI* than English but does not seem to be very different from Dutch or German, contrary to our prediction. Future research using more speakers needs to be done in order to validate the role of the *rPVI* in capturing rhythmic patterning of different languages.

Our findings support the prediction for Japanese. Japanese is mora-timed, and it patterns with the syllable-timed languages. A mora is a linguistic unit that is often smaller than a syllable, but in terms of speech timing, mora-timing is more similar to syllable-timing than to stress-timing. The comparable vocalic *nPVI* values for Japanese and French agree with this observation. Neither language has vowel reduction. In contrast, the intervocalic *rPVI* values for Japanese are in the region of those exhibited by the stress-timed languages English, Dutch and German. Initially, this is surprising, since Japanese has a relatively simple syllable structure. Vowels between voiceless consonants are often devoiced in Japanese, however, and are not associated with formant patterns in a spectrogram. As stated above, we have taken devoiced vowels to be part of an intervocalic, not a vocalic, interval. Potentially devoiced vowels constitute 16% of the vowels in the Japanese North Wind passage (cf. IPA Handbook, 1999). Our approach to measuring intervocalic intervals probably raised the intervocalic *rPVI* value for Japanese.

Figure 2 contains the results for all languages in our corpus. Table 1 in the appendix gives the normalised vocalic and the raw intervocalic values for the complete set of data. Consider first the *PVI* values from the rhythmically mixed languages Polish and Catalan.

Nespor (1990) argued that Polish is rhythmically mixed because the language does not have vowel reduction, but can have very complex syllable-structures. Catalan was said to be mixed because it resembles syllable-timed languages in syllable-structure, but does have vowel reduction, unlike syllable-timed Spanish.

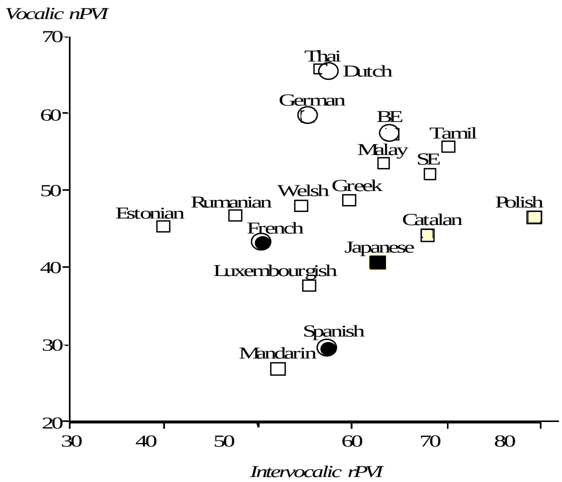

Figure 2. PVI profiles for data from eighteen languages. Prototypical ○ = stress-timed, ● = syllable-timed, ■ = mora-timed, □ = mixed or unclassified

Our data provide an acoustic basis for Nespor’s claims. The vocalic *nPVI* value for Polish is similar to that for syllable-timed French. But on the intervocalic axis, the two languages are some considerable distance apart. In fact, the intervocalic *rPVI* value for Polish is the highest in our set.

Figure 2 also supports for Nespor’s (1990) observations on vocalic differences between Spanish and Catalan. These languages are separated on the vocalic *nPVI* axis. But the vocalic *nPVI* from Catalan, which has vowel reduction, is similar to that obtained from French, which does not. This finding illustrates a point made by Low et al. (2000) who compared spectral patterns of reduced vowels in Singapore English and British English. Significant differences appeared in the way vowels are reduced in these varieties of English. From a phonological point of view, Singapore English has vowel reduction, but reduced vowels are less centralised in the F1/F2 space than reduced vowels in British English. Reduced vowels in Singapore English are also longer than their counterparts in British English. These findings suggest that we may be able to account for our vocalic *nPVI* data from French, Spanish, and Catalan on the basis of differences in vowel quality and vowel reduction.

We will discuss the remaining findings beginning with Thai. This language patterns with the stress-timed group (Dutch, German, British English). Thai was classified as stress-timed by Luangthongkum (1977). Our findings support his view. Singapore English, marked SE in Figure 2, was classified as syllable-timed by Tongue (1979), Platt and Weber (1980), Yeow (1987). Our data show that Singapore English exhibits slightly less vocalic variability than British English. However, Singapore English is not at all close to the traditional syllable-timed languages French or Spanish. Luxembourg and Mandarin pattern with the syllable-timed group. Our Mandarin data provide the lowest vocalic *nPVI* of all languages investigated in the present study. Overlapping with the edges of the stress-timed and the syllable-timed group, we find the unclassified languages Welsh, Greek, Malay, Tamil and Rumanian. Estonian exhibits the lowest intervocalic *rPVI* value. Apparently, with respect to intervocalic variability, Estonian is the opposite of Polish. Finally, the findings for Tamil go against Corder's (1973) and Asher's (1985) classification of Tamil as syllable-timed. We found high vocalic *nPVI* and high intervocalic *rPVI* values for Tamil.

3.2 *Further analyses*

Since we have data from only one speaker per language, we investigated the stability of *PVI* values within each speaker. We split the vocalic and intervocalic values from each of our eighteen speakers into three equal subsections. This procedure supplied three pairs of values from each speaker. The data are given in Table 2 in the appendix.

Then we carried out statistical analyses on the data. Since we have data from only one subject per language, the results of the analyses are preliminary. Firstly, we tested whether the values from the three vocalic and the three intervocalic sections correlate across languages. All correlations were highly significant (vocalic intervals: sections 1 and 2: r=0.926, p<0.01, sections 1 and 3: r=0.844, p<0.01, sections 2 and 3: r=0.788, p<0.01; intervocalic intervals: sections 1 and 2: r=0.673, p<0.01, sections 1 and 3: r=0.816, p<0.01, sections 2 and 3: r=0.624, p<0.01, 2-tailed). The stability of our measurements seems quite satisfactory.

Secondly, we performed an analysis of variance on the data (SPSS, General Linear Model, repeated measures) with the dependent variables vocalic *nPVI* and intervocalic *rPVI* and the within-subjects factor Section (1,3). Between–subjects effects were highly significant, but the factor Section was not (vocalic *nPVI* F[1,17]= 729.6, p<0.001, intervocalic rPVI F[1,17]= 347.8, p<0.001; NB: our between subject-effects are equivalent to between-language effects). Then we carried out post hoc tests on the data (SPSS, General Linear Model, multivariate analysis, Tukey).

Table 2 shows which of the vocalic *nPVI* and the intervocalc *rPVI* differences between languages were significant (Tukey, p<0.05).

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

Table 2. Post hoc multiple comparisons for rPVI values, x = significant vocalic nPVI difference, p<0.05; o = significant intervocalic rPVI difference, p<0.05.

In Table 2, the upper half of the matrix shows the results for the vocalic *nPVI*. The lower half shows the results for the intervocalic *rPVI*. The number of significant differences for the *nPVI* measure was twice as large as the number of differences for the *rPVI* measure. Apparently, the *nPVI* provides a better separation of languages than the *rPVI*. Table 2 shows that the *nPVI* values from the prototypical stress-timed languages German, English and Dutch differ from the *nPVI* values from with the prototypical syllable-timed languages French and Spanish. But the languages whose values are located between those from the prototypical stress-timed and the prototypical syllable-timed groupare not significantly different from either (see Figure 2). Fewer differences emerged for the intervocalic measure: basically, Polish and Estonian differ from each other and from most of the other languages in the sample. The so-called rhythmically mixed language Catalan differs from Spanish on the vocalic but not on the intervocalic axis. Polish, also classified as mixed, is significantly different from most other languages on the intervocalic axis. The vocalic *nPVI* values from Polish differ from the German and Dutch *nPVI* values, but not from the British English values.

Finally, we compared our measures with those developed by Ramus et al. for rhythmic classification. Ramus and colleagues suggested that the best measure of rhythmic diversity is provided by *%V*, the proportion of time in an utterance devoted to vowels, and **Δ** , the standard deviation of intervocalic intervals. We calculated these measures from our data. The results appear in Figure 3. We have plotted *%V* on the y-axis in descending order, to allow a direct comparison with our vocalic *nPVI* findings. Table 3 in the appendix gives the *%V* and **Δ** values.

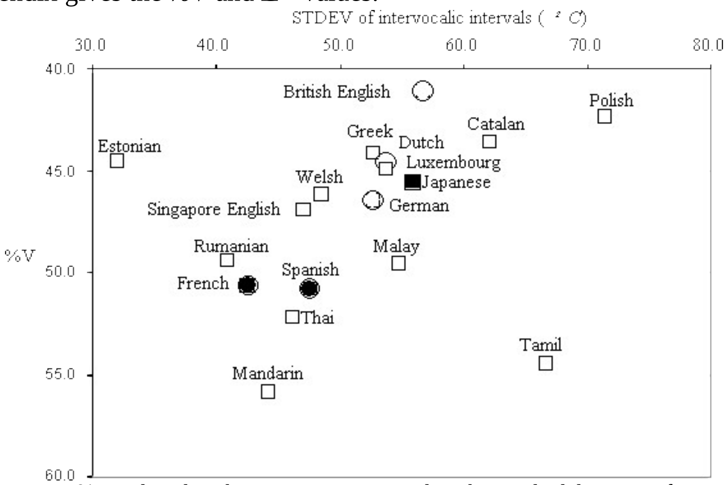

Figure 3. The measure *%V* is plotted on the y-axis, in reverse order. The standard deviation of intervocalic intervals **Δ** , is given on the x-axis.

Figure 3 shows that the **Δ***%V/* measure provides results similar to those provided by the combined *PVI* for Estonian, Polish, and Mandarin. The *rPVI* places Estonian and Polish at the extremes of the intervocalic axis, and so does the **Δ** . Mandarin is the language with the highest *%V* value, and it has the lowest vocalic *nPVI*.

The results for German, Dutch and English are similar on both sets of measures. These languages exhibit relatively low *%V* values, and mid-range **Δ** values. But Greek, Catalan, Welsh, and Luxembourg appear in the stress-timed area in Figure 3. They exhibit lower *%V* values than German and appear to be more stress-timed than German.

Two languages move from the stress-timed group in Figure 2 to the syllable-timed group in Figure 3. Thai, which is supposedly stress-timed and has the highest vocalic *nPVI*, moves into the syllable-timed area. Tamil, another language with a high vocalic *nPVI*, also moves there. Corder (1973) and Asher (1985) suggested that Tamil is syllable-timed, but Thai was said to be stress-timed (Luangthongkum, 1977). The results for Thai and Tamil show that languages which exhibit a high proportion of vowels in utterances may also exhibit a high level of variability in vocalic intervals. Note that this is not the case for English (high vocalic *nPVI*, low *%V*) or Spanish (low vocalic *nPVI*, high *%V*).

The %V measure places Japanese between German and Dutch, within the stress-timed group. Polish exhibits a %V value between those of German and British English. This finding does not support Nespor's suggestion that Polish patterns with syllable-timed languages with respect to vowels.

4. Discussion

4.1 *Rhythm classes: categorical or gradient?*

On the vocalic axis, the prototypical stress-timed languages German, English and Dutch are well separated from the syllable-timed languages French and Spanish. On the basis of this result, one could offer a categorical distinction between stress- and syllable-timing (for want of better terms for the rhythmic groupings in question). But the data also show that languages can be more or less stress-timed or syllable-timed. Therefore, a strict categorical distinction between stress-timing and syllable-timing cannot be defended. In some areas of linguistics, strict categorical distinctions exist, for instance, in syntax. Either a word is a member of a syntactic category, or it is not. The nature of the rhythm class distinction is different: we find degrees of stress- or syllable-timing. Distinctions which are similar in nature to the one between rhythmic groupings in our data are found in speech perception. The studies by Samuel (1977) and by Carney, Widen and Viemeister (1977) and colleagues showed that although certain speech continua were perceived categorically by naïve subjects, the labelling function became less distinct with practice while discrimination performance improved. Categorically perceived continua are not necessarily perceived in an absolute or discrete manner (Harnad, 1987). We will refer to this effect as 'weak categorical'. Our data show that there is a weak categorical distinction between the group of languages that has been described as stress-timed, and the group of languages that have been described as syllable-timed.

Secondly, the results show that there is overlap between the stress-timed and the syllable-timed group and unclassified languages, and that Japanese is not in a rhythm class of its own. Therefore, although we find a weak categorical distinction between stress- and syllable-timing, it is clear that not all languages of the world fit into that distinction.

Thirdly, the results show that the vocalic *nPVI* separates languages into a stress-timed and a syllable-timed group, but the intervocalic *rPVI* does not. Instead, the intervocalic *rPVI* shows why Polish does not fit into either of the prototypical rhythm classes, and why Estonian may be difficult to classify. Polish is different because the intervocalic *rPVI* is very high. Estonian is different because the *rPVI* is very low.

Finally, the data show that languages that exhibit an extreme level of durational variability in one dimension have non-extreme variability in the other. The plot in Figure 2 has a distinct diamond shape. A definition of rhythm in speech as the recurrence of similar acoustic units at relatively regular intervals may account for this finding. If vocalic as well as intervocalic intervals were extremely variable, and independent, then there would be no recurrence of similar acoustic units, and hence no impression of rhythmicity.

4.2 *Measures of rhythm class: The vocalic PVI and %V*

We have found comparable results for the extremes of the *PVI* space and the *%V/* space proposed by Ramus, Nespor and Mehler (1999). The locations of Estonian, Polish, Mandarin and British English are similar in both spaces. But in the centre of the space, we find differences. The *%V* values show that the proportion of vowel time in Greek, Catalan, Welsh, Luxembourg and Japanese is lower than in German. Accordingly, these languages should be more stress-timed than German. Our data contradict this assertion: the vocalic *nPVI* is lower in Greek, Catalan, Welsh, Luxembourg and Japanese than in German. Hence, the *nPVI* suggests that these languages are less, not more stress-timed than German.

The switch of Tamil and Thai from the stress-timed group to the syllable-timed group was particularly noticeable under the *%V/* measure. Thai is associated with a very high vocalic *nPVI*, but *%V* also is high. The same observation holds for Tamil, although vocalic *nPVI* values for Thai are more extreme. The relationship between *%V* and the vocalic *nPVI* in traditional stress- or syllable-timed languages is different. In British English and Spanish *%V* and vocalic *PVI* values seem complementary. Figure 4 illustrates our point. The figure shows that British English and German have low *%V* values and high vocalic *nPVI* values. French and Spanish have high *%V* values, but low vocalic *PVI* values. In Thai and Tamil, *%V* values are higher than in French and Spanish. But unlike in French and Spanish, the vocalic *nPVI* is high also.

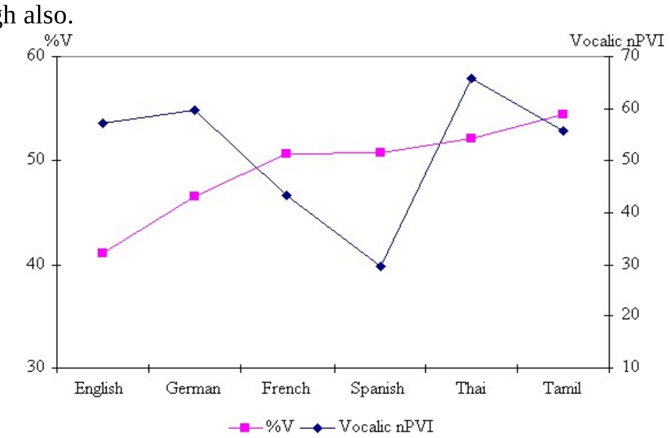

Figure 4. Left y-axis: *%V*; right y-axis: vocalic *nPVI* values. The variety of English is British English.

This complementarity of overall vowel time and vocalic variability in English and German on the one hand, and French and Spanish on the other may contribute substantially to impressions of stress- or syllable-timing. If the relationship between the two measures provides the acoustic basis for an impression of stress- or syllable-timing, then Thai would be classified as stress-timed. Although *%V* is high, the vocalic *nPVI* is even higher. But Tamil would not be classifiable.

5. Conclusion

We have provided acoustic evidence for rhythmic diversity among languages from duration measurements. Unlike other researchers in the field of speech timing, we did not measure interstress intervals or syllable durations which are phonological units. Instead, we took a direct route from impressionistic observations of rhythmic differences between languages to the acoustic signal. We measured the durations of vowels, and the duration of intervals between vowels in a passage of speech. Then we computed an acoustic variability index which expresses the level of variability in vocalic and intervocalic intervals. Our data support a weak categorical distinction between stress-timing and syllable-timing. But the distinction does not encompass all of the world's languages. There is considerable overlap between the stress-timed and the syllable-timed group and hitherto unclassified languages.

Acknowledgements

For comments and suggestions, we are grateful to John Coleman, Anders Eriksson, Brechtje Post, Burton Rosner, Andrew Slater, Ian Watson and Briony Williams. We would also like to thank an unnamed reviewer and the editors for their help. For translations and assistance with recruiting subjects we would like to thank Eva Liina Asu, Eva Estebas-Vilaplana, Eric Fixmer, Keiko Masuda, Elinor Reynolds, Catherine Sangster, Nancy Waugh and Paula West. Finally, we are grateful to Andrew Slater for IT assistance. The research reported here was supported by the U.K. Economic and Social Research Council (ESRC) through grant no. R000237145.

Endnotes

[1] The contribution of vowel reduction to the impression of stress- and syllable-timing was investigated more generally in studies by Wenk and Wioland (1982) and by Brakel (1985). These authors suggested that the basis for stress- or syllable-timing may rest on vowels rather than syllables. In British English, there is the presence of vowel reduction and the effect of this is that the duration of each foot is nearly isochronous. Syllable-timed languages like French do not have reduced vowels and as such, do not seem to achieve foot isochrony.

[2] For an investigation of rhythm which also departs from isochrony, but in a different direction, see Cummins and Port (1998). Cummins and Port define rhythm in speech as the hierarchical organization of temporally coordinated prosodic units.

[3] We did not apply this index to a mora-timed language.

[4] A further consequence of our approach to measuring in the present paper is that we do not include the hold phase of intonation-phrase-initial stops after a silence interval.

[5] Note that we could have normalised by dividing by the average interval duration. We have retained the normalisation procedure suggested by Deterding (1994) as it is more sensitive to local changes in speech rate. Local changes may be especially relevant in languages like French, which, crudely, have rhythmic structures consisting of very similar intervals within rhythmic groups, with considerable phrase-final lengthening at the end of each group. For a prosodic analysis of French intonation, see Post, 2000.

[6] A further reason not to apply the normalisation procedure for vocalic intervals to intervocalic intervals arises from research on speech rate effects on vowels and consonants. Gay (1978) has shown that increases in speaking rate in English lead to a shortening of both consonantal and vocalic portions of syllables, but most of the change results from a shortening of the vocalic portions. Changes in the duration of initial and final formant transitions in CVC sylllables accounted for about one third of the total shortening.

6. Appendix

Languages	Normalised Vocalic *nPVI*	N	Raw Intervocalic *rPVI*	N
Thai	65.8	161	56.5	164
Dutch	65.5	132	57.4	136
German	59.7	155	55.3	153
BE	57.2	124	64.1	124
Tamil	55.8	149	70.2	150
Malay	53.6	205	63.3	204
SE	52.3	118	68.2	118
Greek	48.7	177	59.6	179
Welsh	48.2	152	54.7	150
Rumanian	46.9	183	47.6	182
Polish	46.6	124	79.1	128
Estonian	45.4	162	40.0	158
Catalan	44.6	144	67.8	139
French	43.5	146	50.4	142
Japanese	40.9	176	62.5	177
Luxembourg	37.7	131	55.4	139
Spanish	29.7	173	57.7	156
Mandarin	27.0	141	52.0	135

Table 1. Normalised vocalic *nPVI* and intervocalic *rPVI* values. The table is sorted in ascending order by vocalic *nPVI* values.

	Vocalic n*PVI*	Intervocalic r*PVI*	N
Thai	52.6	69.5	53
	63.0	71.1	53
	53.3	56.4	53
Dutch	60.6	66.6	44
	41.8	70.8	44
	55.2	59.4	44
German	52.1	57.6	51
	57.0	65.3	51
	55.9	58.7	51
BE	65.6	55.2	40
	65.0	53.6	40
	54.4	56.1	40
Tamil	70.1	56.1	50
	67.8	53.9	50
	72.8	56.4	50
Malay	55.1	53.1	68
	60.4	60.0	68
	63.0	48.2	68
SE	70.6	49.9	39
	69.8	45.3	39
	64.5	58.7	39
Greek	57.2	51.5	59
	58.0	43.9	59
	61.5	54.7	59
Welsh	50.9	43.6	50
	50.6	48.2	50
	59.2	46.5	50
Rumanian	42.1	52.2	60
	47.4	46.1	60
	49.7	39.0	60
Polish	71.6	48.6	42
	77.8	46.3	42
	80.3	43.2	42
Estonian	37.0	49.6	53
	39.0	49.4	53
	38.7	41.1	53
Catalan	66.0	46.0	47
	52.8	47.9	47
	62.1	38.5	47
French	49.3	39.4	46
	49.7	38.7	46
	44.3	42.0	46
Japanese	56.3	39.9	58
	71.3	42.9	58
	47.0	40.3	58
Luxembourg	52.3	30.1	43
	58.2	39.4	43
	54.0	37.5	43
Spanish	60.3	30.5	52
	56.9	28.0	52
	54.7	31.2	52
Mandarin	52.0	26.4	45
	55.0	27.7	45
	44.2	26.0	45

Table 2. Vocalic *nPVI* and intervocalic *rPVI* data subdivided into three sections. Sorted as Table 1. above to allow for comparisons, i.e. by mean vocalic *nPVI* in ascending order.

	%V	%V Ramus et al.	ΔV	ΔV Ramus et al.	ΔC	ΔC Ramus et al.
British English	41.1	40.1	46.6	46.4	56.7	53.5
Polish	42.3	41.0	44.9	25.1	71.4	51.4
Catalan	43.6	45.6	33.9	36.8	62.1	45.2
Greek	44.1		49.1		52.7	
Estonian	44.5		39.6		31.9	
Luxembourg	44.7		31.1		53.7	
Dutch	44.9	42.3	48.4	42.3	53.7	53.3
Japanese	45.5	53.1	53.0	40.2	55.8	35.6
Welsh	46.1		39.4		48.5	
German	46.4		44.5		52.6	
Singapore English	46.9		41.0		47.0	
Rumanian	49.4		49.5		40.9	
Malay	49.5		56.7		54.8	
French	50.6	43.6	35.5	37.8	42.4	43.9
Spanish	50.8	43.8	20.7	33.2	47.5	47.4
Thai	52.2		74.8		46.1	
Tamil	54.4		76.4		66.6	
Mandarin	55.8		36.2		44.1	

Table 3. Comparison of our findings with data from Ramus et al. (1999). NB. Ramus and Colleagues investigated Italian, but Italian was not investigated in the present paper.

8. References

Abercrombie, D. (1965) Studies in Phonetics and Linguistics. London: Oxford University Press.

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

Asher, R.E. (1985) Tamil. London: Croom Helm.

Beckman, M.E. (1982) Segment duration and the 'mora' in Japanese. Phonetica 39:113-135.

Beckman. M.E. (1992). Evidence for speech rhythms across languages. In Y. Tohura, E. Vatikiotis-Bateson and Y. Sagisaka (editors), Speech Perception, Production and Linguistic Structure, pp. 457-63, Tokyo: Omsha and Amsterdam: IOS Press.

Beckman, M.E. & Edwards, J. (1994) Articulatory evidence for differentiating stress categories. In Phonological structure and phonetic form. Papers in Laboratory Phonology III. (P. Keating, editor), pp. 7-33, Cambridge: Cambridge University Press.

Bloch, B. (1950). Studies in colloquial Japanese IV: Phonemics. Language 26, 86-125.

Bolinger, D.L. (1965) Forms of English: Accent, Morpheme, Order. Cambridge, Massachusetts: Harvard University Press.

Brakel, A. (1985) Towards a morphophonological approach to the study of linguistic rhythm, Chicago Linguistic Society, 21, pp. 15-25.

Carney, A.E., Widen, B. & Viehmeister, N. (1977) Non-categorical perception of stop consonants differing in VOT, Journal of the Acoustical Society of America, 62, pp. 961-970.

Catford, J. C. (1977) Fundamental Problems in Phonetics. Edinburgh: Edinburgh University Press.

Classe, A. (1939) The Rhythm of English Prose. Oxford: Blackwell.

Corder, S. P. (1973) Introducing Applied Linguistics. Harmondsworth, Middlesex: Penguin.

Couper-Kuhlen, E. (1990) Discovering rhythm in conversational English: Perceptual and acoustic approaches to the analysis of isochrony, KontRI Working Paper No. 13, University of Konstance, Fachgruppe Sprachwissenschaft.

Couper-Kuhlen, E. (1993) English Speech Rhythm. Form and Function in Everyday Verbal Interaction. Amsterdam: Benjamins.

Cummins, F. and Port, R. F. (1998). Rhythmic constraints on stress timing in English. Journal of Phonetics, 26(2):145-171.

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

Dauer, R.M. (1983) Stress-timing and syllable-timing re-analysed, Journal of Phonetics 11, pp. 51-62.

Dauer, R. (1987) Phonetic and phonological components of language rhythm, Proceedings of the XIth International Congress of Phonetic Sciences, Talinn, Estonia, pp. 447-450.

Delattre, P. (1966) A comparison of syllable length conditioning among languages. International Review of Applied Linguistics in Language Teaching, IV(3), pp. 183-198.

Deterding, D.H. (1994) The rhythm of Singapore English. Paper presented at the Fifth Australian International Conference on Speech Science and Technology. December 6-8, 1994.

Eriksson, A. (1991) Aspects of Swedish speech rhythm. Gothenburg Monographs in Linguistics, 9, University of Göteborg.

Faure, G. Hirst, D.J. and Chafcouloff, M. (1980) Rhythm in English: Isochronism, pitch, and perceived stress. In The Melody of Language (L.R. Waugh and C.H. van Schooneveld, editors), pp. 71-79, Baltimore: University Park Press.

Fischer-Jørgensen, E. & Hutters, B. (1981) Aspirated stop consonants before low vowels, a problem of delimitation, its causes and consequences ARIPUC, 15, pp. 77-102.

Gay, T. (1978) Effects of speaking rate on vowel formant transitions. Journal of the Acoustical Society of America, 63, pp. 223-230.

Grabe, E., Post, B. & Watson, I. (1999) The acquisition of rhythm in English and French, Proceedings of the Intonational Congress of Phonetic Sciences, 2, pp. 1201-1204. San Francisco, August 1999

Handbook of the International Phonetic Association. (1999) Cambridge: Cambridge University Press.

Harnad, S. (1987) Categorical Perception. Cambridge: Cambridge University Press.

Han, M.S. (1962) The feature of duration in Japanese. Onsei no kenkyuu 10, pp. 65-80.

Hockett, C.F. (1958) A Course in Modern Linguistics. New York: Macmillan.

Hoequist, C.J. (1983a) Durational correlates of linguistic rhythm categories. Phonetica, 40, pp. 19-31.

Hoequist, C.J. (1983b) Syllable duration in stress-, syllable- and mora-timed languages. Phonetica 40, pp. 203-237.

Kohler, K. (1982) Rhythmus im Deutschen, Arbeitsberichte, Institut für Phonetik der Universität Kiel, 19, pp. 89-106.

Ladefoged, P. (1967) Three Areas of Experimental Phonetics. London: Oxford University Press.

Ladefoged, P. (1975) A Course in Phonetics. New York: Harcourt Brace Jovanovich.

Laver, J. (1994) Principles of Phonetics. Cambridge: Cambridge University Press.

Lehiste, I. (1977) Isochrony reconsidered. Journal of Phonetics, 5, pp. 253-263.

Low, E.L. (1998) Prosodic prominence in Singapore English. Unpublished Ph.D. Thesis, University of Cambridge.

Low, E.L. & Grabe, E. (1995) Prosodic patterns in Singapore English, Proceedings of the Intonational Congress of Phonetic Sciences, 3, pp. 636-639, Stockholm 13-19 August.

Low, E.L., Grabe, E. & Nolan, F. (2000) Quantitative characterisations of speech rhythm: 'syllable-timing' in Singapore English. Language and Speech, in press, to appear in autumn 2000.

Luangthongkum, T. (1977) Rhythm in Standard Thai. Unpublished Ph.D. thesis, University of Edinburgh.

Manrique, A.M.B. & Signorini, A. (1983) Segmental reduction in Spanish. Journal of Phonetics, 11, pp. 117-128.

Nakatani, L.H., O'Connor, J.D. & Aston, C.H. (1981) Prosodic aspects of American English speech rhythm, Phonetica, 38, pp. 84-105.

Nespor, I. (1990) On the rhythm parameter in phonology. In Logical issues in Language Acquisition (I. Roca, editor), pp. 157-175, Dordrecht: Foris.

Oyakawa, T. (1971). On the directionality of segmental conditioning in Japanese. Monthly Internal Memo, Department of Linguistics, University of California at Berkely, pp. 81-103.

Peterson, G. E. & Lehiste, I. (1960) Duration of syllable nuclei in English, Journal of the Acoustical Society of America, 32, pp. 693-703.

Pike, K. (1946) The Intonation of American English. 2nd edition. Ann Arbor: University of Michigan Press.

Platt, J.T. & Weber, H. (1980) English in Singapore and Malaysia: Status, Features and Functions. Kuala Lumpur: Oxford University Press.

Pointon, G.E. (1980) Is Spanish really syllable-timed? Journal of Phonetics 8, 293-304.

Port, R.F., Al Ani, S. & Maeda, S. (1980) Temporal compensation and universal phonetics. Phonetica, 37, pp. 235-252.

Port, R.F., Dalby, J. and O'Dell, M. (1987) Evidence for mora-timing in Japanese. Journal of the Acoustical Society of America 81, pp. 1574-85.

Post, B. (2000) Tonal and Phrasal Structures in French Intonation. Thesus: The Hague.

Ramus, F., Nespor, M. & Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal, Cognition, 72, pp. 1-28.

Roach, P. (1982) On the distinction between 'stress-timed' and 'syllable-timed' languages. In Linguistic Controversies (D. Crystal, Editor), pp. 73-79, London: Arnold.,

Samuel, A.G. (1977) The effect of discrimination training on speech perception: noncategorical perception, Perception & Psychophysics, 22, pp. 321-330.

Shen, Y. & Peterson, G.G. (1962) Isochronism in English, University of Buffalo Studies in Linguistics, Occasional Papers, 9, pp. 1-36.

Smith, A. 1976. Syllable-structure and rhythm in Japanese. Work in Progress, Department of Linguistics, University of Edinburgh, 8, pp. 1-8.

Tongue, R.K. (1979) The English of Singapore and Malaysia. Singapore: Eastern Universities Press.

Wenk, B. & Wioland, F. (1982) Is French really syllable-timed? Journal of Phonetics, 10, pp. 193-216.

Yeow, K.L. (1987) Stress, Rhythm and Intonation in Educated Singapore English. Unpublished Masters Thesis, National University of Singapore.