Journal of Speech Sciences 1(2): 3-15. 2011
Available at: http://www.journalofspeechsciences.org

EXPLORING THE SPEECH RHYTHM CONTINUUM: EVIDENCE FROM ASHANTI TWI

OLIVO, Ann Marie

Rice University

In current speech rhythm research, the traditional search for isochrony and speech rhythm classes has been replaced with a focus on uncovering the acoustic correlates of rhythm in the speech signal (Ramus, Nespor, and Mehler 1999, Grabe and Low 2002, Cummins 2002). In this paper, I present findings from a study in which I describe speech rhythm in a language whose rhythm had not been previously studied—Ashanti Twi (Niger-Congo, Kwa). Additionally, I test the validity of claims made about the utility of various rhythm metrics.

Two native speakers of Ashanti Twi participated in the study. Each speaker was recorded while reading a translation of "The North Wind and the Sun". Vocalic and consonantal intervals were measured in Praat using auditory and visual cues. Various rhythm metrics (interval measurements and Pairwise Variability Indices) were then computed and compared to results from prior studies.

Results show that Ashanti Twi is rhythmically more similar to languages that have been traditionally described as syllable-timed, such as French and Spanish. However, it does not fall clearly into the traditional stress- or syllable-timed categories, supporting the claim that speech rhythm should be studied as a continuum rather than a categorical distinction.

Keywords: phonetics; speech rhythm; rhythm metrics; Akan; Ashanti Twi

*Author for correspondence: amolivo@gmail.com

1 Introduction

Current trends in speech rhythm research have focused on uncovering the acoustic correlates of rhythm in the speech signal. This is a departure from the traditional search for isochrony characteristic of earlier work that was concerned with categorizing languages into either syllable- or stress-timed rhythmic patterns. In this paper, I present findings from a study conducted with two goals in mind. The first was to describe speech rhythm in a language whose rhythm pattern had not been studied, Ashanti Twi (Niger-Congo, Kwa). The second goal of the study was to test the validity of claims made about the utility of various rhythm metrics. In Section 1, I will review the literature on rhythm metrics. In Section 2, I will provide a brief overview of Akan phonology. In Sections 3 and 4, I will describe the methodology and results of the study. The paper concludes with a discussion in Section 5.

1.1 Background

The languages of the world have traditionally been classified into two rhythm classes: stress-timed or syllable-timed (c.f. Pike 1945, Abercrombie 1967). This categorical distinction was initially based on researchers' own perceptions of stress and beats as well as a timing assumption that was based on chest-pulses—the amount and intensity of air flowing from the lungs during speech: "Speech rhythm is essentially a muscular rhythm, and the muscles concerned are the breathing muscles" (Abercrombie 1967:96).

According to the traditional definition, syllable-timed languages are those in which all syllables are said to be isochronous (such as French, Yoruba, and Telugu) and stress-timed languages are those in which the durations between the stressed syllables are said to be isochronous (such as English, Russian, and Arabic) (Abercrombie 1967). Furthermore, the distinction was considered categorical. Abercrombie wrote:

> "There are two basically different ways in which the chest-pulses and the stress-pulses can be combined, and these give rise to two main kinds of speech-rhythm. As far as is known, every language in the world is spoken with one kind of rhythm or with the other" (1967:97).

Pike (1945) is often credited with establishing this classification system for the types of rhythm found across languages, but he was actually referring to stress-timed vs. syllable-timed *rhythm units*. These units were utterances "spoken with a single rush of syllables uninterrupted by a pause" (Pike 1945:34). In fact, he never claimed that languages should fall into one of either category, nor did he claim the "isochronous" units would be exactly equal (Kohler 2009). Pike asserted that English demonstrated both stress-timed and syllable-timed *rhythmic units*, but that the syllable-timed unit is used only rarely in certain speech styles. He described the two speaking styles in English that use syllable-timed rhythmic units as "somewhat" *staccato* (where unstressed syllables are each pronounced abruptly with a measured beat on each syllable) and *spoken chant* (where unstressed syllables are "more or less" equally timed and glided together). No example was provided for the staccato style. Childhood teasing in English was given as an example of spoken chant as in "Susie is a tattle tale" (34). It should also be noted that Pike proposed this distinction as a way to describe intonation patterns in English to Latin American English-language learners. The distinction was not made with languages besides English, Spanish, and Portuguese in mind.

Roach (1982) doubted the plausibility of experimental work concerning speech rhythm since syllable identification relies on phonological theory and stress identification must be done auditorily where the investigator may be biased by his or her own native language. For example, an English-speaking investigator may mistake phonemic vowel lengthening or lexical tone for stress. Dauer (1983) questioned the necessity of the labels of "stress-timed" and "syllable-timed" for languages. She suggested researchers use the term "stress-based" to talk about how important stress is in relation to the perception of speech rhythm in different languages, and also that there are both phonetic (e.g. vowel reduction) and phonological (e.g. syllable structure) phenomena that can contribute to the perception of speech rhythm. Dauer (1987) proposed a system of rating phonetic (length, pitch, quality) and phonological (syllable structure, function of accent) components of languages in order to more accurately describe the rhythmic types found cross-linguistically. Dimitrova (1998) used this method to describe speech rhythm in Bulgarian and found that it placed Bulgarian between prototypically stress- and syllable-timed languages. She concluded that the methodology could be used as a starting point for further investigation. As phonetic investigation has advanced in recent years, this methodology is not utilized in any current research, but the research agenda that drove it is still vital.

Cummins (2002) provided an argument against the categorization of speech rhythm into classes by means of a speech cycling task in which he measured the P-centers of vowels, which correspond to the vowels' onsets. He claimed that different types of rhythm are achieved by means of hierarchical organization between different levels of linguistic and prosodic units, i.e. syllables, feet, phrases, etc. He suggested that rhythm is "hung" onto prosodic units, but that these linguistic units can be forcibly entrained, since the interaction is between discrete (speech sounds) and gradient units (rhythm). Additionally, the degree of entrainment within speakers and across utterances varies greatly. Cummins concluded that, "a gradient phenomenon, not yet well understood, mediates the role of syllables in determining macroscopic timing patterns. Its gradient nature precludes it from supporting a classification among languages" (2002:2). As was pointed out by Kohler (2009), Cummins' view is influenced greatly by the work of Classe (1939), which was the first (and often-overlooked) attempt at measuring isochrony in speech rhythm by means of a kymograph[1]. Classe was also the first scholar to suggest that speech rhythm is influenced by grammatical and contextual features. According to this model, three factors are integrated to produce English rhythm "(a) A phonetic factor, the number of units in the bar. (b) A logical factor, the grammatical connexion between the bars. (c) A phonetic factor, the nature of the units in the bar, especially accents" (Classe 1939:89).

Another current approach uses a dynamic coupled oscillator model to describe speech timing (cf. e.g. Barbosa 2002, Barbosa, Viana, and Trancoso 2009, O'Dell and Nieminen 1999, 2009). An oscillator refers to a process that repeats itself regularly. When this process is influenced by another process, this is referred to as coupling. In this way, speech signal can be affected by articulatory gestures, for example (O'Dell and Nieminen 2009). Timing in this framework is described by the concept of *phases* (the relationship of time to oscillators). Furthermore, this framework has the advantage of being applied to other rhythmic phenomena, not just speech sounds.

2 Rhythm Metrics

As researchers distanced themselves from the search for isochrony and moved towards finding more precise ways of comparing perceptually different kinds of speech rhythm, rhythm metrics

[1] A kymograph was an instrument used to measure the temporal variation in muscular contractions (i.e. blood pressure) by means of a rotating drum that produced graphical representations of the differences in pressure.

were developed as a way to acoustically measure the phonetic correlates of speech rhythm (as opposed to phonological correlates such as lexical stress and syllable structure). Ramus, Nespor, and Mehler (1999) developed a set of interval measurements in order to test their hypothesis about how infants perceive contrasting rhythms. Based on the findings of Nazzi, Bertoncini, and Mehler (1998) that newborns pay more attention to vocalic intervals rather than consonants, they hypothesized that infants should perceive language as a series of vocalic and consonantal intervals. They developed three measurements to extract rhythm from the speech signal: the proportion of vocalic intervals (%V); the standard deviations of consonantal intervals (ΔC); and the standard deviation of vocalic intervals (ΔV). They then plotted the data based on (%V, ΔC), (%V, ΔV), and (ΔV, ΔC) planes. By means of an ANOVA, they found that the (%V, ΔC) projection corresponded most clearly to the rhythm classes established by intuition and perception ($p < 0.001$). They found that traditionally categorized stress-timed languages, which typically have more complex syllable structures, will have a higher ΔC. These languages tend to also exhibit a lower %V, although this may not be true for languages that exhibit vowel length distinctions.

Grabe and Low (2002) measured durational variability in successive acoustic intervals using the Pairwise Variability Index (PVI). The equation for the raw Pairwise Variability Index (rPVI) is presented in Equation 1 below, where *m* is the number of intervals and *d* is the duration of the k^{th} interval. This measure is used for the consonantal[2] intervals in order to capture differences in phonotactics. A normalized version of this equation is used for the vocalic intervals. The equation for the normalized Pairwise Variability Index (nPVI) is presented in Equation 2 below where *m* is the number of intervals and *d* is the duration of the k^{th} interval. The normalized equation is used for vocalic intervals in order to control for speech rate. They used these measurements on data from 18 languages, including 7 languages previously unclassified per rhythm type. They concluded that their data provided evidence for a weak categorical distinction, but that this distinction does not account for all languages of the world.

Equation 1: The raw Pairwise Variability Index

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

Equation 2: The normalized Pairwise Variability Index

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

Dellwo (2006) calculated the coefficient of variation of consonantal and vocalic segments—referred to as VarcoC and VarcoV, respectively—in order to normalize the interval measurements of Ramus, Nespor, and Mehler (1999). In languages with more complex syllables, consonant clusters may get reduced as speech rate increases (Dellwo 2006). The Varco measurements are meant to normalize that variability. The measurement was compared to the interval measurements across five different speech rates in English, German, and French. Dellwo claimed that VarcoC is a more reliable measurement than ΔC because it shows less within-language variation across different speech rates. Roach (1998), however, pointed out that

[2] Grabe and Low (2002) use the term 'intervocalic' instead of 'consonantal'. I prefer the term 'consonantal' as it corresponds to the terminology of the other rhythm measurements. Furthermore, pauses between segments were omitted in the analysis, leaving purely consonantal segments and not intervocalic segments.

since speech rate varies greatly within-languages, it may not be meaningful to compare speech rates cross-linguistically. Additionally, other researchers have shown that speech rate can differ for the same language depending on whether the measurement was taken to be speech sounds per second or syllables per second. For example, den Os (1988) found that when measured in syllables per second, Italian and Dutch were similar in speech rate; however, when measured in speech sounds per second, Italian had a slower speech rate.

White and Mattys (2007) tested all the above-mentioned rhythm metrics in order to compare how well they captured the rhythm class continuum. Their results patterned with the original results obtained by Ramus, Nespor, and Mehler (1999) in regard to defining broad rhythm classes, but the correlations between speech rate and the metrics ΔC and ΔV suggested the need for normalization. They found that the normalized metric for ΔC, VarcoC, showed no discrimination, which they attribute to the normalization process, while the normalized metric for ΔV, VarcoV, did discriminate between rhythm classes at varying speech rates. They also found that the nPVI-V and rPVI-C scores did not reflect the rhythm classes distinguished in Grabe and Low (2002). They found the (%V, VarcoV) projection to be most insightful as it captured both variations in speech rate as well as vocalic segmental normalization.

Loukina et al. (2009) evaluated the success of 15 rhythm metrics in distinguishing languages of different rhythm types. They found that not all rhythm metrics were equally successful cross-linguistically, but overall the vocalic measurements were more successful than the consonantal measurements. They also found that differences in speech rate did not have an effect on perceived differences between languages.

Additional studies have attempted to use different rhythm metrics to analyze previously unstudied languages. Gut et al. (2001) studied rhythm in three West African tone languages in the Niger-Congo language family—Anyi (Kwa), Ega (most commonly categorized as Kwa, although some researchers have disputed this), and Ibibio (Benue-Congo). They used the metrics developed by Ramus, Nespor, and Mehler (1999) as well as the Rhythm Index (Low and Grabe 1995), the precursor to Grabe and Low's PVI (2002). Furthermore, they compared the classification to syllable structure in the languages to test Dauer's (1983) claim that syllable structure and vowel length would be reflected in the rhythm class (i.e. stress-timed languages should exhibit more complex syllable structure). Their results suggest that all three languages fall into the syllable-timed category. They found no evidence that syllable structure or vowel lengthening had an affect on rhythm classification.

Gut and Milde (2002) compared the rhythm of Nigerian English to British English as well as the languages studied in Gut et al. (2001). Their results indicated that Nigerian English falls mid-way between British English and the other West African languages studied. They also found that Nigerian English had a greater percentage of vowel intervals than British English, as well as a higher percentage of open syllables. These results establish Nigerian English as a language that cannot be clearly categorized according to the classical distinction, suggesting the need to view the different types of speech rhythm as a continuum rather than a categorical distinction.

In summary, current trends in the field have led researchers to abandon the pursuit of isochrony in exchange for developing methods to more accurately describe the perceptual differences in rhythm types. Among the various rhythm metrics that have been developed, it has been claimed that (%V, VarcoV) is the superior metric for comparing rhythm types (White and Mattys 2007). Applying these methodologies to previously unstudied languages provides insight into the finer distinctions of speech rhythm types. The present study ventures to apply these methodologies to Ashanti Twi in order to describe its speech rhythm in terms of a continuum rather than attempting to place it within the categorical class distinction.

3 Ashanti Twi

Ashanti Twi is considered a dialect of Akan [ISO aka] that is mutually intelligible with Fante Twi and Akuapem Twi. It is a register tone language of the Kwa branch of the Niger-Congo language family spoken by 2.8 million people in Ghana (Lewis 2009).

3.1 Phonology

According to the current literature, Twi[3] allows three syllable structures: V, C, and CV (Ofori 2006, Marfo and Yankson 2008). The data in (1) come from Ofori's (2006) work on hypocorisms in Twi.

(1)	(i)	V	/**a.ku.a**/	'Wednesday-born female'
	(ii)	CV	/**su**/	'cry'
	(iii)	C	/**n.ko.su.a**/	'eggs'

For the current study, a fourth syllable type is included in the analysis: CCV. An example of this type is /**a.kro.ma**/ 'vulture'. The only consonant that occurs as the second consonant in the cluster is /r/. Marfo and Yankson (2008) provided a formal analysis of the CrV syllable structure in Twi, claiming it is only realized in a phonetic form and that the underlying structure is /CV.CV/. Their analysis is based on Dolphyne's (1988) book on Akan phonology where she claimed that words in which the CrV syllable structure occurs may be pronounced with a high vowel inserted between the consonant and /r/ in careful pronunciations. Since we are interested in a phonetic analysis of speech rhythm in Twi, the complex onset structure is included in the analysis as it occurs in the data.

The syllable is considered the tone-bearing unit in Twi. Twi has two level tones (High and Low) that interact in various ways to convey both lexical and grammatical differences (Dolphyne 1988).

To date, no claims have been made as to speech rhythm in Twi. Because it allows for an occasional complex onset and exhibits vowel devoicing, I hypothesized that Twi would fall towards the stress-timed end of the continuum. However, after debriefing one of the consultants following the experiment, she indicated that she thought Twi "felt" more like French in terms of syllable structure, putting it on the syllable-timed end of the spectrum. Additionally, closely related languages have been described as syllable-timed (Gut et al. 2001). Based on these observations, Twi could present a case where prototypical properties of stress-timed languages co-occur with features of prototypical syllable-timed languages, reiterating the need to move towards a gradient classification system for rhythm across languages, as has been suggested by Dauer (1983, 1987) and others.

4 Method

4.1 Participants

Two speakers participated in the study. Speaker 1 is a 27-year-old female graduate student at Rice University. She is a native speaker of Ashanti Twi and started learning English at age 6. She moved from Accra to the United States at age 18, where she has lived for the past 9 years, with the exception of one year she spent studying in England. She has no formal instruction in

[3] From this point on 'Twi' will refer to Ashanti Twi and neither of the other dialects.

Ashanti Twi. Her mother is ethnically Ashanti and her father is Ewe. She also speaks French and Gã.

Speaker 2 is a 27-year-old male graduate student at Rice University. He is a native speaker of Ashanti Twi who started learning English at age 6. He moved to the United States from Accra 9 years ago. He received formal instruction in Ashanti Twi in school. Both his parents are ethnically Ashanti. He also speaks Gã.

4.2 Data

The data were collected using "The North Wind and the Sun", which is a standard text in phonetic research. Since a Twi translation of the text was not available, the text used was translated and transcribed by Speaker 1 and the author as per Grabe and Low (2002). Speaker 1 was recorded in a sound-proof booth at Rice University in Houston, TX and Speaker 2 was recorded in a quiet classroom, also at Rice University. Both recordings were done using lapel microphones on a Marantz PMD 660 at a sampling rate of 44 kHz. Each speaker read the text three times. For each speaker, the version that was read most clearly with the fewest interruptions was the one used for the analysis.

4.3 Procedure

The sound files were segmented in Praat into vocalic and consonantal intervals using auditory and visual cues based on prior work on speech rhythm and acoustic analysis (cf. Peterson and Lehiste 1960). Vocalic intervals only consisted of the signal in which there were visible formants between the vowel onset and offset. In other words, the vowel was taken to begin at the start of the second formant; likewise, it ended at the end of the second formant. In instances of vowel devoicing, the vowel was considered to be consonantal rather than vocalic as per Grabe and Low (2002) in order to base the analysis solely on phonetic segments[4]. In instances of aspirated stop-vowel pairs, the aspiration was considered part of the consonantal segment as per White and Mattys (2007). Pre-vocalic glides were segmented as consonantal and post-vocalic glides were segmented as vocalic as per Ramus, Nespor, and Mehler (1999). Instances of silence between words or sentences were omitted.

Once the files were segmented into a TextGrid, a Praat script was run to calculate the proportion of vocalic intervals (%V), the standard deviations of consonantal intervals (ΔC), and the standard deviation of vocalic intervals (ΔV). The script also produced the rate-normalized interval measures VarcoC and VarcoV. The intervocalic and consonantal durations calculated by the script were then put into an Excel worksheet to calculate the nPVI-V and rPVI-C values. The data were then run through the projections suggested by White and Mattys (2007) and compared to their results.

5 Results

The interval measurements ΔV, ΔC, %V, VarcoV, VarcoC, and speech rate collected from the Praat script are presented in Table 1 below along with the PVI scores calculated by the Excel worksheet.

[4] This would be problematic in a language like Japanese, which is hypothesized to base its timing on the phonological unit of the mora (e.g. Port, Dalby, and O'Dell 1987). Japanese exhibits high vowel devoicing and the decision to count voiceless vowels as consonantal would incorrectly place Japanese among the stress-timed languages.

Table 1: Rhythm metrics for Speaker 1 and Speaker 2 of Twi

	Speaker 1	**Speaker 2**
Interval Measures		
ΔV	46.76	48.51
ΔC	69.38	72.95
%V	42.25	43.74
VarcoV	50.84	50.66
VarcoC	50.28	55.07
Pairwise Variability Indices		
nPVI-V	48.87	44.47
rPVI-C	72.32	75.98
Speech Rate		
Syllables/second	4.64	4.59

For cross-linguistic comparison, the data from Table 1 above have been rounded and are appended with the first language data from White and Mattys (2007). These data are presented in Table 2 below.

Table 2: Rhythm metrics for Twi compared to the measurements from White and Mattys (2007). Spanish and French are intended to be prototypical examples of syllable-timed languages while English and Dutch are prototypical examples of stress-timed languages.

	Spanish	**French**	**Twi 1**	**Twi 2**	**English**	**Dutch**
Interval Measures						
ΔV	32	44	47	49	49	49
ΔC	40	51	69	73	59	49
%V	48	45	42	44	38	41
VarcoV	41	50	51	51	64	65
VarcoC	46	44	50	55	47	44
Pairwise Variability Indices						
nPVI-V	36	50	49	45	73	82
rPVI-C	43	56	75	76	70	52
Speech Rate						
Syllables/second	8.0	5.6	4.6	4.6	5.2	6.0

Figure 1 below shows the distribution of languages over the (%V, ΔC) plane. It can be seen that both measurements for Twi exhibit a high ΔC. This was partly expected because Twi exhibits both a complex onset and vowel devoicing. However, the ΔC measurement is greater than both English and Dutch, which allow for much more complex syllables. It can be deduced that Twi exhibits such a high ΔC because voiceless vowels were analyzed as consonantal segments, which led to the high variability of consonantal segments. This projection does not place Twi within either rhythm class. It falls in between the two classes according to the %V measurement.

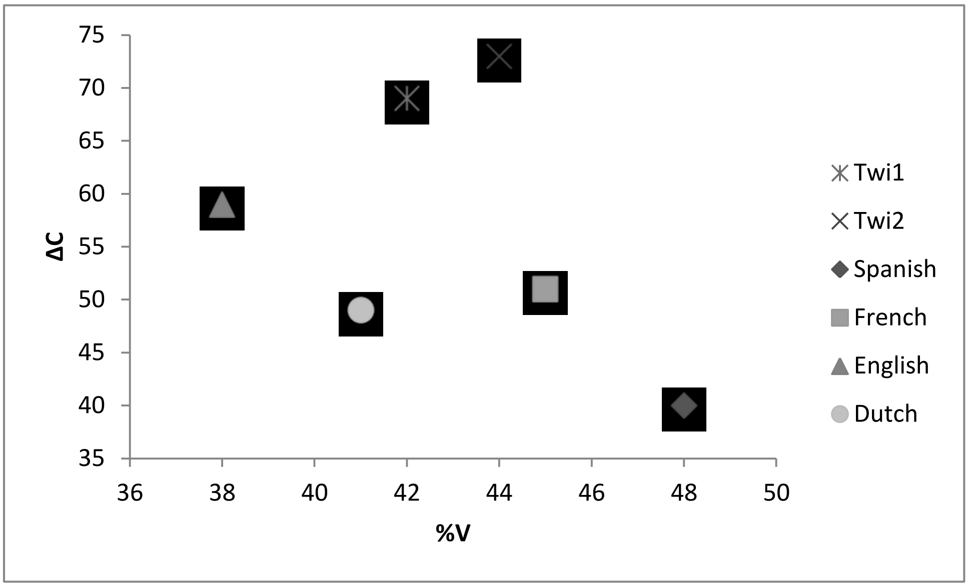

Figure 1: Distribution of languages over the (%V, ΔC) plane.

Figure 2 below shows the distribution of languages over the (rPVI-C, nPVI-V) plane. Again, because of the syllable structure and voiceless vowels found in Twi, it has a high rPVI-C. It is not clearly categorized in either rhythm class, but it appears to align more with the syllable-timed languages with regard to its nPVI-V score.

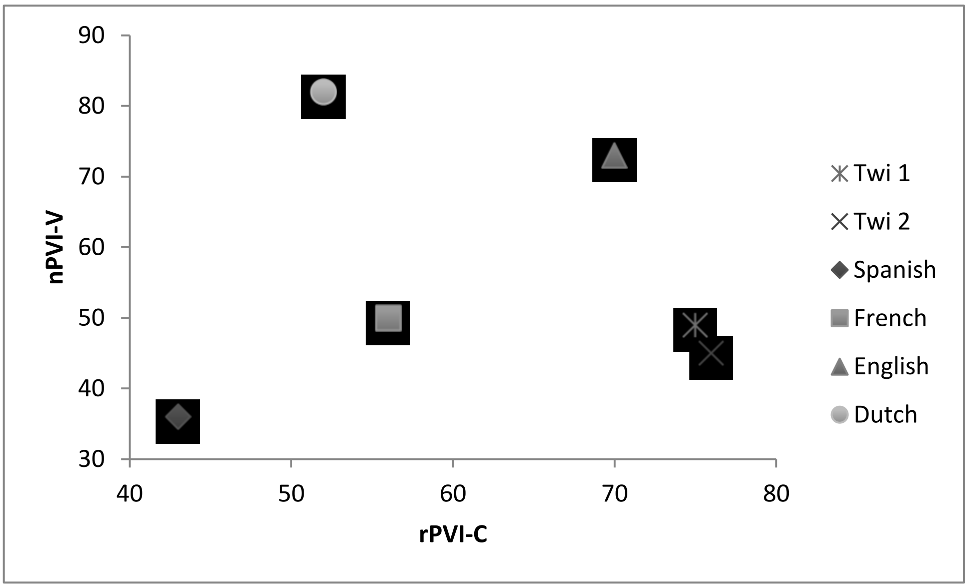

Figure 2: Distribution of languages over the (rPVI-C, nPVI-V) plane.

Figure 3 below shows the distribution of languages over the (%V, VarcoV) plane. White and Mattys (2007) found this projection to align most closely with the traditional groupings. This projection places Twi between the syllable-timed and stress-timed languages, suggesting it is nearly as syllable-timed as French, but not as much as Spanish.

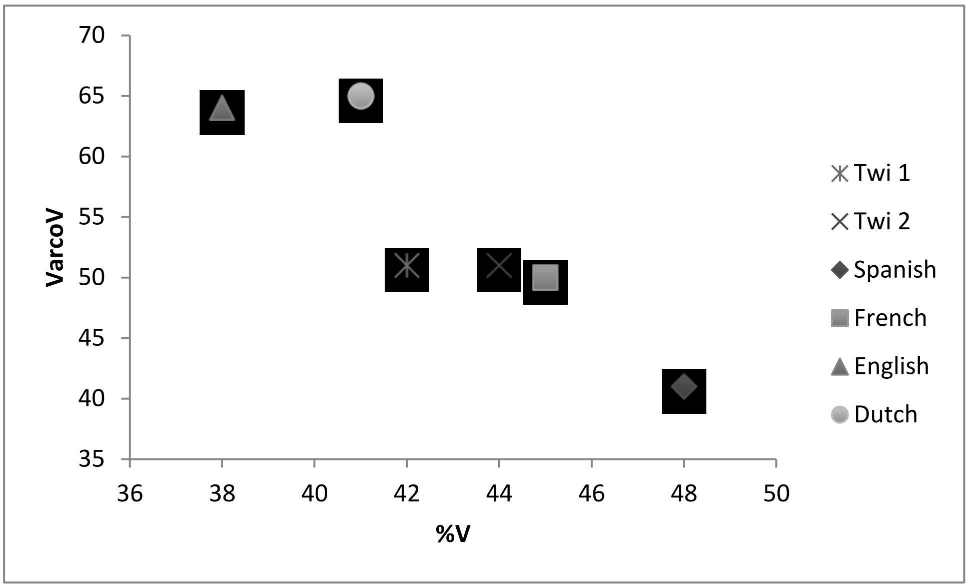

Figure 3: Distribution of languages over the (%V, VarcoV) plane.

6 Discussion

Compared to the results of White and Mattys (2007), the interval measurements and PVI scores for Twi do not clearly categorize it with either the stress-timed languages or syllable-timed languages, supporting the expectation that Twi would not be classified into either of the two categories. By using the normalized version of the interval measurement ΔV (Varco V), the continuum of languages by rhythm type was captured, with Twi falling closer to the syllable-timed end.

The high ΔC and rPVI-C values are demonstrative of consonantal variation due to the vowel devoicing and the complex onset that occur in Twi. The projections over the (%V, ΔC) and (rPVI-C, nPVI-V) planes did not demonstrate categorical results. Ramus, Nespor, and Mehler (1999) and White and Mattys (2007) demonstrated that consonantal measurements were only effective in distinguishing rhythm classes when speech rate was kept constant. The speech rate between the two Twi speakers did not vary as is shown in Table 1 above, but it was different from the speech rates of the speakers of Spanish, French, English, and Dutch from the White and Mattys (2007) study, as is shown in Table 2 above. This could also account for why the distributions involving consonantal measurements did not show clear results. The ΔC measurement was meant to correct for speech rate, but according to the current data, it was unsuccessful. All these factors suggest the superiority of VarcoV as a reliable and practical rhythm metric.

Grabe and Low (2002) found that languages with high nPVI-V and a high rPVI-C are more stress-timed, and languages with the opposite results would be more syllable-timed. Twi has a low nPVI-V and a high rPVI-C indicating it falls in between the categorization of stress-timed and syllable-timed. Likewise, Twi exhibited a fairly high %V and a high ΔC indicating it falls in between the two rhythm classes.

The projection on the (%V, VarcoV) plane suggests that Twi is more syllable-timed. Interestingly, White and Mattys (2007) concluded that the (%V, VarcoV) projection both reinforced the typology of rhythm classes and captured the small variation within classes. The Twi data support this idea. Although Twi falls towards the traditionally syllable-timed languages, the rhythm varies slightly between the two speakers. Within the category, it is also shown that Twi is comparably as syllable-timed as French, but not as syllable-timed as Spanish, demonstrating the need to classify languages as "more or less" stress- or syllable-timed as pointed out by Dauer (1983) and others.

7 Acknowledgements

Research for this project was carried out during a Field Methods course at Rice University from August 2008 to May 2009. Funding for the consultants was provided by the Rice University Dolores Mitchell Fund. I would like to thank the two consultants on this project, Dzifa Duose and Ben Appiah for their time and effort, and especially Dzifa for her help in the translation of the North Wind and the Sun text. I would also like to thank Katherine Crosswhite, Nancy Niedzielski, Christina Willis, Laura Robinson, and two anonymous reviewers for their insightful and constructive feedback on earlier drafts of this paper. The Praat script used to measure the durations was written by Tae-Jin Yoon of University of Victoria. The Excel worksheet used to calculate the PVI values was modified from the example on Esther Grabe's website: www.phon.ox.ac.uk/˜esther/ to include the equation for the rPVI. Remaining inaccuracies are of course my own.

REFERENCES

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

Barbosa, Plínio, 2002. Explaining cross-linguistic rhythmic variability via a coupled-oscillator model of rhythm production. In *Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence. 163–166.

Barbosa, Plínio, M. Céu Viana, and Isabel Trancoso, 2009. Cross-variety rhythm typology in Portuguese. In *INTERSPEECH-2009*. Brighton. 1011–1014.

Classe, Andre. 1939. *The rhythm of English prose*. Oxford: Blackwell.

Cummins, Fred. 2002. Speech Rhythm and Rhythmic Taxonomy. In *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence, France.

Cummins, Fred and Robert F. Port. 1998. Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26. 145–171.

Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11. 51-62.

Dauer, Rebecca M. 1987. Phonetic and Phonological Components of Language Rhythm. In *Proceedings of the Eleventh International Congress of Phonetic Sciences*. 447-450. Tillinn, Estonia.

Dellwo, Volker. 2006. Rhythm and speech rate: A variation coefficient for ΔC. In Karnowski, P. and I. Szigeti (Eds.), *Language and language processing: Proceedings of the 38th linguistic colloquium*. 231-241. Frankfurt: Peter Lang.

den Os, E. 1988. *Rhythm and Tempo of Dutch and Italian,* Utrecht: Drukkerij Elinkwijk.

Dimitrova, Snezhina. 1998. Bulgarian Speech Rhythm: Stress-Timed or Syllable-Timed? *Journal of the International Phonetic Association* 27. 27-33.

Dolphyne, Florence A. 1988. *The Akan (Twi-Fante) Language: Its Sound System and Tonal Structure*. Accra: Ghana Universities Press.

Grabe, Esther and Ee Ling Low. 2002. Durational variability and speech rhythm. In Gussenhoven, C. and Warner (eds.), *Laboratory Phonology 7.* N. Berlin: Mouton de Gruyter.

Gut, Ulrike, Eno-Abasi Urua, Sandrine Adouakou and Dafydd Gibbon. 2001. Rhythm in West African tone languages: a study of Ibibio, Anyi and Ega. In *Proceedings of the Typology of African Prosodic Systems Workshop*. Bielefeld University, Germany.

Gut, Ulrike and J.T. Milde. 2002. The prosody of Nigerian English. In B. Bel and I. Marlien (eds.) In *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence, France.

Kohler, Klaus J. 2009. Rhythm in speech and Language: A New Research Paradigm. *Phonetica* 66. 29-45.

Lewis, M. Paul (ed.), 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.

Loukina, Anastassia, Greg Kochanski, Chilin Shih, Elinor Keane, and Ian Watson. 2009. Rhythm measures with language-independent segmentation. In *INTERSPEECH 2009*. Brighton. 1531-1534.

Low, Ee Ling and Esther Grabe. 1995. Prosodic patterns in Singapore English. *Proceedings of the XIIIth International Congress of Phonetic Sciences* 3. Stockholm. 636-639.

Marfo, Charles and Solace Yankson. 2008. The Structure of the CCV Syllable of Akan. *Concentric: Studies in Linguistics* 34(2). 85-100.

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

O'Dell, Michael and Tommi Nieminen, 2009. Coupled oscillator model for speech timing: Overview and examples. In: Vainio, M., Aulanko, R., Aaltonen, O. (Eds.), In *Nordic Prosody: Proceedings of the Xth Conference*, Helsinki 2008.

O'Dell, Michael and Tommi Nieminen, 1999. Coupled oscillator model of speech rhythm. In: Ohala, J., Hasegawa, Y., Ohala, M., Granville, D., Bailey, A. (Eds.), In *Proceedings of the XIVth International Congress of Phonetic Sciences*. Vol. 2. University of California, Berkeley. 1075–1078.

Ofori, Seth. 2006. Akan Hypocorisms: A Constraint-Based Approach. In Obeng, Samuel Gyasi and Cecilis Sem Obeng (eds.). *From Linguistics to Cultural Anthropology: Aspects of Language, Culture and Family Issues in Ghana (West Africa)*. Meunchen: LINCOM GmbH. 67-86.

Peterson, Gordon E. and Ilse Lehiste. 1960. Duration of Syllable Nuclei in English. The Journal of the Acoustical Society of America 32(6). 693-703.

Pike, Kenneth. 1945. *The Intonation of American English.* 2nd edition. Ann Arbor: University of Michigan Press.

Port, Robert F., Jonathan Dalby, and Michael O'Dell. 1987. Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81. 1574-1585.

Ramus, Franck, Marina Nespor, and Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73. 265-292.

Roach, Peter. 1982. "On the distinction between 'stress-timed' and 'syllable-timed' languages." In Crystal, David (ed.). *Linguistic Controversies*. 73-79.

Roach, Peter 1998 "Some Languages are Spoken More Quickly Than Others" in Language Myths, eds. L. Bauer and P. Trudgill, Penguin, 1998, pp. 150-8

White, Laurence and Sven L. Mattys. 2007. Calibrating rhythm: First language and second language studies. *Journal of Phonetics* 35. 501-522.

Appendix 1: The North Wind and Sun in Ashanti Twi

o.kʷãjn.tu.fwo.bi be.tɕʷa: mu a: o.de n.tu ma: a.tɕɛ.tɕre ne hũ a.bra: na m.fra.ma na: e.fri pje.mu e.nɛ e.ɥja e:.kũ n.tõ.kʷa
A traveler came to pass by who had a cloth around his body while the North Wind and the Sun were arguing.

o.mu tõ.kʷa nʊ fa ni.je ne hũ e dĩ pa: hũ
Their argument was about who was stronger.

o.mu je o.mu a.ʥĩ se ɲe o.be tu.mi a.ma o.kʷãjn.tu.fwo nʊ e.ji n.tu.ma nʊ e.fri ne hũ ka.nje: nʊ o.nu na ne hũ e.dĩ ʨɛ̃ ne joŋ.kũ
They decided that whoever can make the traveler take off his cloth, he's the one who is stronger than his friend.

e.na m.fra.ma na: e.fri pje.mu nʊ hu: m.fra.ma dĩ.dĩ:.dĩ nen sʊ o:.hu nɛ.dɪ nʊ o.kʷãjn.tu.fwo nʊ dĩ ni n.tu.ma na a ʨɛ.ʨrɛ ni hũ dĩ.dĩ:.dĩ a.ko.si bre a: m.fra.ma na: e.fri pje.mu nʊ bre:.je na .oʥa: je
Then the North Wind blew hard, however, the more it blew, the traveler wrapped his cloth around himself tighter until the North Wind got tired and stopped.

e.na e.ɥia nʊ bɔ dĩ.dĩ:.dĩ e.ʌa:n.a o.kʷãjn.tu.fwo nʊ ji n.tu.ma nʊ fri ne hũ
Then the Sun shone down strongly and the traveler immediately removed his cloth.

n.ti m.fra.ma na: e.fri pje.mu nʊ ʥɪ tu mu sɛ e.ɥa na ne hũ e dĩ pa:
And so the North Wind acquiesced that the Sun was the stronger.