

# Experiencing English rhythm through note values\*

Noriko Hattori

**Abstract** This paper describes a method of conveying an appreciation of English rhythm to non-native speakers of English by using musical scores. Comparison of the scores of English, French, and Japanese songs reveals a sharp difference between English on one hand and French and Japanese on the other in the distribution of syllable duration. Counting syllables and plotting histograms of note values is an easy way of experiencing English rhythm especially for learners of English whose native language is not stress-timed. It is a simplified, visual counterpart of the normalized pairwise variability index (nPVI), which calculates durational contrast between successive elements in a sequence.

## 1. Introduction

English is known as a ‘language which is spoken with a stress-timed rhythm’ and ‘its syllables are of uneven length’ (e.g., Jones 1972; Abercrombie 1965; Cruttenden 2008). Some phoneticians have tried to show syllable quantities in English by means of musical notation (e.g., Abercrombie 1965; Allen 1965; Jones 1972; Cruttenden 2008). For instance, the differences of length in phrases (1) and (2) below are expressed by means of musical notation as in (3) and (4), respectively (Abercrombie 1965).

(1) Buy the book/ day to day/ (I) quite forgot

(2) Either book/ (A) later day/ What’s it for?

(3) ♪. ♪♪      (4) ♪♪ ♪

The relationship between speech rhythm and music goes back to an idea of declamation. Declamation is defined as follows: ‘in music, the relation between verbal stress and melodic

accent in the setting and delivery of a text' (Sadie ed. 2001). 'Correct' declamation requires that music respect the rhythmic organization of speech. Although modern vocal music enjoys greater autonomy from the metre of spoken words, a comparison of musical scores of English, French, and Japanese songs reveals a sharp difference in the distribution of syllable duration.

The "normalized pairwise variability index" or nPVI is a quantitative measure of durational contrast between successive elements in a sequence (Low, Grabe & Nolan 2000; Grabe & Low 2002). It is hardly practical for ordinary learners of English to make their own nPVI measurements on speech, since this depends on marking vowel boundaries by segmenting acoustic displays of speech (Patel, Iversen & Rosenberg 2006; Patel 2008), and requires an understanding of spectrograms (Peterson & Lehiste 1960). However, a series of studies have established a parallel between rhythm in speech and music (Ramus, Nespor & Mehler 1999; Patel, Iversen & Rosenberg 2004; Patel, Iversen & Rosenberg 2006; Patel 2008), opening the prospect of using existing song scores as a source of data. The present paper proposes a simple hands-on experience of English rhythm for average Japanese students who are not trained in acoustic phonetics, using song scores as a starting point.

## **2. Methods**

The primary orientation of this study is pedagogical. In English class, Japanese students whose native language is said to be a mora-count syllable-timed language (Dauer 1983) are taught the fact that English is stress-timed. Nonetheless it is doubtful whether average students really feel English rhythm, or to use Abercrombie's term, they are able to 'empathize' (Abercrombie 1965). This study utilizes the idea that "musical notes can roughly be compared to syllables, and vowels form the core of syllables", and "it seems plausible to compare vowel-based rhythmic measures of speech to note-based rhythmic measures of music" (Patel, Iversen & Rosenberg 2006). Each note is supposed to present one syllable. In Figure 1 below, one box, a unit of syllable duration in this study, represents a sixteenth note. Two boxes

represent an eighth note, whose duration is twice that of a sixteenth note. Figure 1 shows a bar of 4/4, which forms a basis of conversion of note values into syllable duration.

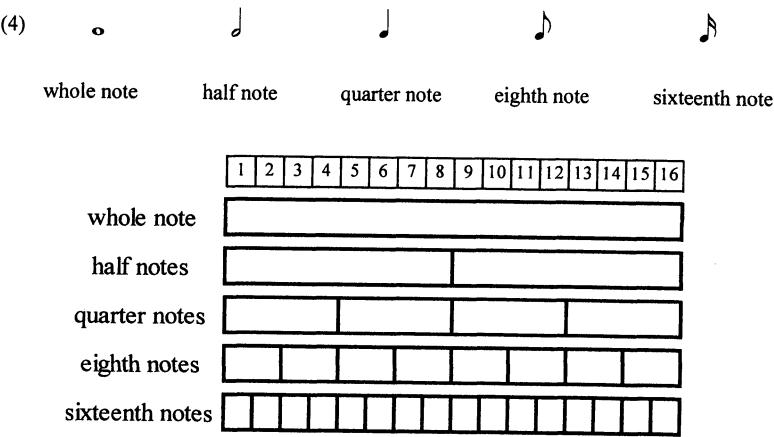


Figure 1: Note values as a basis of calculation of syllable duration

Several songs were selected which were expected to show the anticipated differences clearly. Duration of each note was calculated by examining musical scores and then shown as a series of boxes that correspond to bars in musical notation.

### 3. Results

The result shows that the distribution of syllable duration is not even in English, compared with other two languages. Part of the result of conversion is shown below. Among English songs, *Let it be* and *Michelle* by the Beatles, were chosen to demonstrate the distribution of syllable duration in the lyrics (*The Beatles complete scores* 1993). Figure 2 shows the first twenty bars (with an upbeat start) of *Let it be*. One line, which corresponds to a bar of 4/4, contains 16 boxes each of which then corresponds to a sixteenth note, or a semiquaver. Space between boxes represents rest in a musical score, while (-) means that the preceding vowel is

lengthened.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
															Who	I
A	find	my-	self	in	times					of	rou-	ble				
	Whether	Ma-	ry		comes	to	me								s-	
	put-	ing	words	of	wis-	dom						let	it	be		
		(-)													And	
	in	my	hour	of	dark-	ness	She	is								
	most	ing	right		in	front	of	me							s-	
	put-	ing	words	of	wis-	dom						let	it	be		
B		(-)										Let	it	be		
		(-)		let	it		be					let	it	be		
		(-)		let	it		be									
	Wha-	per	words	of	wis-	dom						let	it	be		
		(-)													And	
C	when	the	bro-	ken	heart-	ed	peo-	ple								
	liv-	ing	in	the	world	e-	gree									
	There	will	be	an	an-	swer						let	it	be		
			(-)												For	
	though	they	may-	be		part-	ed					then	is			
	still	a	chance		that	they	will	see								
	There	will	be	(an)	an-	swer						let	it	be		

Figure 2: Conversion of note values into syllable duration in the first twenty bars of *Let it be*  
Figure 3 shows a distribution of note values in Figure 2.

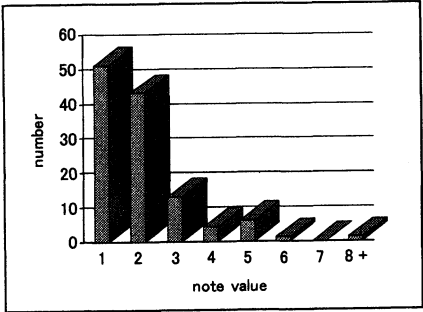


Figure 3: Distribution of note values in Figure 2

By contrast, *Michelle* contains a different, “un-English” syllable distribution, which can be explained as a deliberate choice made by the songwriter. Figure 4 shows the first twenty-two bars of the song. According to the lyrics, this song is about a French lady friend, *Michelle*, whose comprehension of English is rather limited (*‘I love you — Until I find a way, I will say the only words I know that you’ll understand’*). The bars A-3 and A-4 sung in English correspond to those in B-3 and B-4 sung in French (see Figure 4). A note accompanying the manuscript lyric *Michelle* at the British Library says that Paul McCartney attempted to “write a French-sounding song” (The Sir John Ritblat Gallery: Treasures of the British Library).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16												
A	Mi-					chelle																						
	ma					belle																						
	These		are		words		that																					
	go		to-		geth-		er																					
	well				my				Mi-		chelle																	
	( )																											
B	Mi-					chelle																						
	ma					belle																						
	sont		des		mots		qui																					
	vont		tres		bien		en-																					
	semble				tres (3/8)				bien (3/8)		en-(3/8)																	
	semble														I													
C	love (3/8)		you (3/8)		I (3/8)		love (3/8)		you (3/8)		I (3/8)																	
	love		you																									
	that		all		I		(1) (3/8)		want (3/8)		you (3/8)																	
																	say											
	un-		till		I		find		a		we-																	
	( )																I		will									
	say		the		on-		ly																					
	words		I		know		that																					
	you'll				un-				der																			
	stand																											

Figure 4: Conversion of note values into syllable duration in the first twenty-two bars of *Michelle*

Figure 5 shows a distribution of note values in Figure 4.

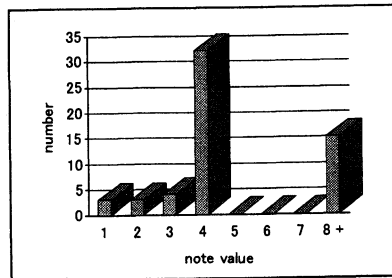
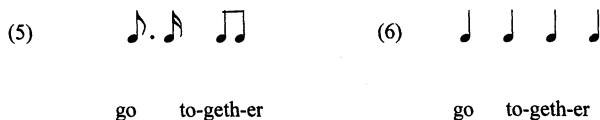


Figure 5: Distribution of note values in Figure 4

Especially the distribution of syllable duration in A-4 is unlike English. It would require musical notation something like (5), not (6).



A French chanson shows preference for type (6) over (5). Here follows an example of a French chanson (*Sans toi ma mie*).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
						Je	sais	tout	est	fi-					
ra-															
(-)						J'ai	per-	du	ta	con-					
fian-						ce									
						Ne-	an-	moins	je	te					
prie															
(-)						de	m'a	ccor-	der	ma					
chan-						ce									
						Si	de-	vant	mon	re-					
mord															
(-)						tu	restes	in	dif-	fe-					
ren-						te									

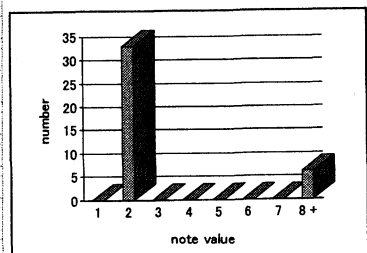


Figure 7: Distribution of note values in Figure 6

**Figure 6: Conversion of note values into syllable duration in the first twelve bars of a French chanson (*Sans toi ma mie*)**

An example of a Japanese song (*Sakurazaka*) follows.

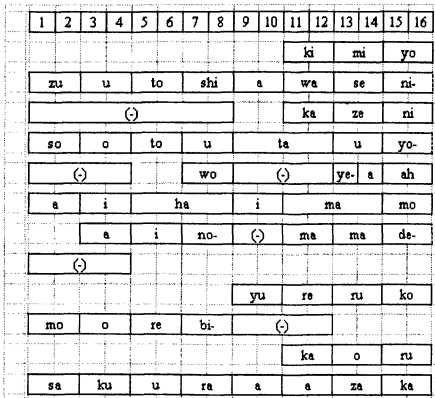


Figure 8: Conversion of note values into syllable (mora) duration in the first twelve bars of a Japanese song (*Sakurazaka*)

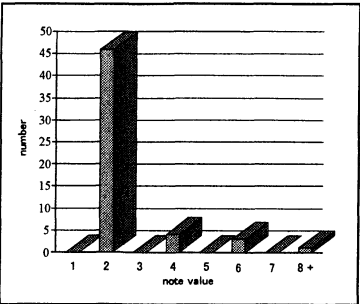


Figure 9: Distribution of note values in Figure 8

It is clear from these figures that the variety and arrangement of syllable duration varies greatly between English on one hand and French and Japanese on the other. Visualization of syllable duration in terms of note values is a quick and easy hands-on experience of various rhythms of languages. Sequential distribution of successive note values is shown below: Figure 10 for *Let it be*; Figure 11 for *Michelle*; Figure 12 for a French chanson; Figure 13 for a Japanese song. High values such as the 6th note in Figure 12 and the 11th note in Figure 13 come from lengthened vowels immediately before rest in a musical score.

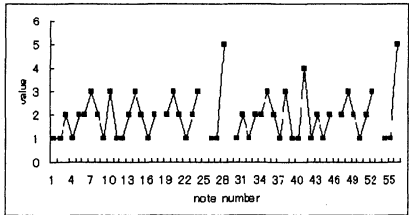


Figure 10: Successive note values for **A** part (the first 56 notes) of *Let it be*

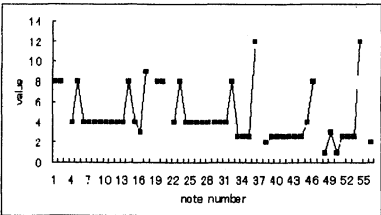


Figure 11: Successive note values for the first 56 notes of *Michelle*

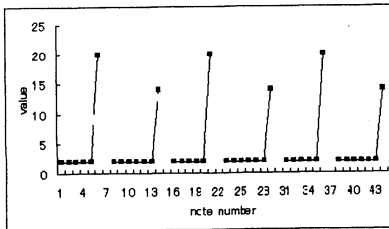


Figure 12: Successive note values for a French chanson (*Sans toi ma mie*)

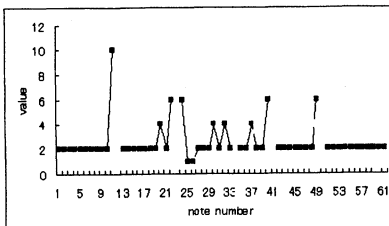


Figure 13: Successive note values for a Japanese song (*Sakurazaka*)

A different distribution of syllable duration in English and Japanese songs is one finding. Another is a different degree of constraint on syllable duration. That is, in the case of Japanese, the duration of each syllable can take a musical note of any duration. Figure 14 shows examples of *sakura* (*sa-ku-ra*, three-syllable (mora) word meaning 'cherry blossom') which appear in three musical scores. In Figure 14, a separate box is given to a lengthened portion of a vowel when the lengthened portion takes a note of different pitch from that of a preceding vowel.

(i)	sa			ku			ra			(-)
(ii)	(a)	sa			ku			ra		
		sa			ku			ra		
	(c)	sa			ku			ra		
					(triple)					
(iii)	sa			ku			u			ra
										a

Figure 14: Lengthening of moras in Japanese songs

This kind of lengthening is limited to the last syllable of a word immediately before rest in the case of English songs. Compared with English, Japanese songs are less constrained in the relative duration of successive syllables, but more constrained in the relative height of successive syllables. Traditional songs and children's songs tend to keep tight declamation. Musical notes reflect prosodic aspects of each language.

#### 4. Discussion



Modern vocal music tends to enjoy autonomy of musical rhythm, and it sometimes ignores inherent characteristics of spoken language deliberately. In spite of this, syllable count in scores is a simple and useful way of demonstrating English rhythm. The Japanese students who attended the author's class and did the syllable count as shown in this paper expressed their interest in the uneven distribution of English syllables. They realized that the function words are given such small space, which explains why they sometimes have trouble comprehending connected English speech. Students can experience different speech rhythm by extending this method to their own and foreign languages. Students can work in a pair or a group of three to help each other in reading musical scores.

Furthermore, the “Englishness” of the above-mentioned songs can be rated by calculating their normalized pairwise variability index (nPVI) scores, whose equation is as follows:

(7)

$$\text{nPVI} = \frac{100}{m-1} \times \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{\frac{d_k + d_{k+1}}{2}} \right|$$

where  $m$  is the number of durations in the sentence (e.g., vowel durations in a sentence) and  $d_k$  is the duration of the  $k$ th element.

The nPVI computes the absolute value of the difference between each successive pair of durations in a sequence, normalized by the mean of these two durations.

Each of these scores ranges between 0 (when the two durations are identical) and 2 (for maximum durational contrast, i.e., when one of the durations approaches zero). The means of these scores, multiplied by 100, yields the nPVI of the sequence (Ramus, Nespor & Mehler 1999; Low, Grabe & Nolan 2000; Patel & Danielle 2003; Patel 2008: 177).<sup>1</sup>

The nPVI scores calculated from the nominal note values for the three songs are shown in Figure 15, which can be compared with those in Figure 16 reported in the previous studies

(Patel, Iversen & Rosenberg 2006: 3040). It is worth noting that their nPVI scores for music are based on instrumental music.

	nPVI	standard deviation	mean
<i>Let it be</i>	58.4	1.008	1.941
<i>Michelle</i>	40.9	2.564	4.553
<i>San toi ma mie</i>	45.6	5.610	4.307

Figure 15: nPVI scores in the three songs

	English nPVI	French nPVI	<i>p</i>
Speech (vowels)	55.0 (3.0)	35.9 (1.8)	< 0.001
Music (notes)	47.1 (1.8)	40.2 (1.9)	< 0.01

Figure 16: nPVI scores in English and French  
(based on Patel, Iversen & Rosenberg 2006: 3040)

The relative position of the three songs in terms of nPVI can be said to reflect the attested difference in nPVI scores in the two languages, which in turn reflects different distribution of syllable duration in them.

### 5. Conclusions

This paper proposes a quick and easy method to experiencing English rhythm by converting note values into syllable duration. It utilizes the fact that one syllable whose core is a vowel is associated with one musical note in scores. Distribution of syllable duration varies between English on one hand and French and Japanese on the other, as pointed out in studies using nPVI. This study is a simplified visual counterpart of nPVI, and ordinary learners of English can easily experience English rhythm using this method.

\*I would like to thank Michael Ashby (UCL) for his useful suggestions on an earlier version of the draft. Any errors are of course my own responsibility. This work was supported in part by a Grant-in-Aid for Scientific Research (C)(2) (Japan Society for the Promotion of Science, Grant No. 2152056).

#### Note

<sup>1</sup>npVI scores can be calculated online by accessing to the following site: [http://vesicle.nsi.edu/users/patel/npvi\\_calculator.html](http://vesicle.nsi.edu/users/patel/npvi_calculator.html)

#### References

- Abercrombie, David. 1965. *Studies in phonetics and linguistics*. London: Oxford University Press.
- Allen, William. S. 1965. *Living English speech*. New edition. London: Longmans.
- The Beatles complete scores*. 1993. Milwaukee, WI.: Hal Leonard Corporation.
- Cruttenden, Alan. 2008. *Gimson's pronunciation of English*. 7th edn. London: Arnold.
- Dauer R.M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11. 51-62.
- Grabe, Esther & Low, Ee L. 2002. Durational variability in speech and the rhythm class hypothesis. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology* 7. 515-546. Berlin: Mouton de Gruyter.
- Jones, Daniel. 1972. *An outline of English phonetics*. Cambridge: Cambridge University Press.
- Low Ee L., Grabe, Esther & Nolan, F. 2000. Quantitative characterization of speech rhythm: Syllable-timing in Singapore English. *Language and Speech* 43. 277-401.
- Patel, Aniruddh D. 2008. *Music, language, and the brain*. Oxford: Oxford University Press.
- Patel, Aniruddh D. & Daniele, J. 2003. An empirical comparison of rhythm in language and music. *Cognition* 87. B35-B45.
- Patel, Aniruddh D., Iversen, J. R. & Rosenberg, J. C. 2004. English and French classical music reflect the melody and rhythm of speech in the two cultures. Popular version of paper 5aSC13 presented at 148<sup>th</sup> ASA Meeting, San Diego, CA.
- Patel, Aniruddh D., Iversen, J. R. & Rosenberg, J.C. 2006. Comparing the rhythm and melody of speech and music: The case of British English and French. *Journal of the Acoustical Society of America* 119. 3034-3047.
- Peterson, G. E. & Lehiste, I. (1960) Duration of syllable nuclei in English. *Journal of the Acoustical Society of*

*America* 32. 693-703.

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

1-28.

Sadie, Stanley (ed). 2001. *The new GROVE Dictionary of music and musicians*. London: Macmillan Publishers.