

Accent and Beat Matching: The Correspondence of English Stress and Japanese Pitch in Terms of Textsetting

Noriko Hattori

Mie University
hattori@human.mie-u.ac.jp

ABSTRACT

This study investigates the relation between speech accent and musical beat in two languages of different accent types. ‘Accent’ is intended to refer to any type of prominence: stress in English and pitch in Japanese. Previous studies on English textsetting indicate a rather strong matching of linguistic accents and musical beat strength. In L2 acquisition, Japanese learners of English tend to substitute pitch height for the ‘Strong-Weak/Weak-Strong’ relation in English. This study examined whether Japanese High-Low pitch pattern would be equivalent to English trochee, and Japanese Low-High pitch pattern would be equivalent to English iamb in terms of textsetting. Selected vocal music scores were examined; the results suggest that English stress and Japanese pitch are not equivalent in terms of textsetting, while they may be in language learners’ production.

Keywords: rhythm, stress, accent, bar, musical beat.

1. INTRODUCTION

Studies of English rhythm have gained attention since the turn of the century when a series of rhythmic measures such as a normalized Pairwise Variability Index (nPVI; the degree of contrast between successive durations in an utterance) and %V (the proportion of vocalic intervals in an utterance) were proposed [5, 7]. They have opened the way to a comparison between the rhythm of speech and that of music [7, 8, 9]. Textsetting is another speech-music interface [1, 2, 3, 4], where alignment of syllables with the rhythmic grid is examined [2]. The following grid notation is based on previous studies [1, 2, 3, 4].

Figure 1: The grid notation used in textsetting problem

whole note	ES:	x		x	
half notes	S:	x	x	x	x
quarter notes	M:	x	x	x	x
eighth notes	W:	x	x	x	x
sixteenth notes	EW:	x	x	x	x

Rows of *x* depict series of isochronic beats, and the columns indicate the strength of individual beats. ‘ES’, ‘S’, ‘M’, ‘W’, and ‘EW’ stand for Extra Strong,

Strong, Medium, Weak and Extra Weak, respectively [1, 2]. They correspond to whole note (semibreve), half notes (minims), quarter notes (crotchets), eighth notes (quavers), and sixteenth notes (semitaunes), respectively. The grid plays the same role as time-signature and bars in conventional musical notation [1].

This study confines itself to vocal music with the time signature of 4/4 (four quarter notes per bar/measure), since the Syllabic Distribution Algorithm (hereafter SDA) [2], which summarizes a rule-based analysis, deals with 4/4. It assumes the conventions of written music [1, 2]:

- The first note after a bar line is given extra emphasis
- A bar of four is divided into two groups of two, though with the third beat rather less strong than the first [10].

4/4 | J J J J J |
strongest weak strong weak

The SDA is summarized below [2]:

- (a) Map stressed syllables one-to-one, left-to-right onto S positions;

After each iteration:

- (b) Map stressless syllables one-to-one, right-to-left onto the highest grid level able to accommodate them.

Once (a) and (b) have applied as many times as they can:

- (c) Map any remaining stressless syllables one-to-one, left-to-right onto the highest grid level able to accommodate them.

Here follow some of the findings made in previous studies directly related to this study:

- Textsetting is predictable and produced by native speakers of the language in question [1, 2, 3]
- Matching between linguistic accents and musical beat strength is evidenced in English; henceforth referred to as ‘stress-to-beat’ or ‘stress-to-strong’ [1, 2].

English and Japanese differ in terms of rhythmic typology: stress-timed vs. mora-timed languages. The principle of ‘stress-to-beat’ simply does not work in Japanese, which is not a stress-timed language. Japanese learners of English tend to substitute ‘High-Low’ relation for the ‘Strong-

Weak' relation in English; accordingly, the Weak-Strong-Weak relation in, for instance, *computer* [kəm'pu:tər] (x X x; where 'X' represents a primary stress, and 'x's represent unstressed syllables) is observed to be converted into a Low-High-Low relation of pitch patterns in Japanese without any loudness or duration added.

2. BACKGROUND

In Standard Japanese, phrases differ in terms of accent only as regards the place where pitch falls. To illustrate the pitch accent system, let us take an example of pitch accent patterns of nouns in Standard Japanese [6]. Among phrases consisting of four syllables (illustrated below by phrases consisting of three-syllable nouns followed by a one-syllable case marker ‘-ga’), there are only the following four distinct patterns:

- (i) makura ga: HLL (L) ‘pillow’
 (ii) kokoro ga: LHL (L) ‘heart’
 (iii) atama ga: LHH (L) ‘head’
 (iv) sakana ga: LHH (H) ‘fish’

where ‘H’ means a high-pitched syllable and ‘L’ means a low-pitched syllable. These four patterns fall into two types: either starting with a high pitched syllable as in (i) or starting with a low pitched syllable as in (ii), (iii), and (iv).

The hypothesis made here is that the first type would be equivalent to English ‘Strong-Weak’ relation, that is, trochee ($X x$), while the second type would be equivalent to English ‘Weak-Strong’ relation, that is iamb ($x X$), if English stress and Japanese pitch are totally equivalent in terms of textsetting. Especially, the second type typically starts as ‘upbeat’ (where a weak beat groups with the following stronger beat [10]) in English vocal music to avoid a mismatch of stress and musical beat. This paper focuses on two aspects of textsetting in Japanese vocal music, and examines whether the following equation will work in textsetting:

- Is Japanese HL pattern equivalent to English trochee (Strong-Weak; X x)?
 - Is Japanese LH pattern equivalent to English iamb (Weak-Strong; x X)?

3. METHOD

An examination of vocal music scores was conducted that include nouns of the two different types of accent pattern. As a representative of the first type of pitch pattern, *sekai* ‘the world’ was chosen, and for the second type, *sakura* ‘the cherry tree/blossoms’ was chosen. How they are set to tunes was examined by consulting the music scores available on the website called @ Elise [11]. Lyrics

that include the nouns in question were searched on the site and then their scores were examined. Due to copyright protection, the whole scores are not shown. We had to disregard candidate songs if their scores did not include the necessary bars where the nouns in question appeared.

4. RESULTS

Here follow how these nouns are set to musical beats in Japanese vocal music. Efforts were made to obtain ten examples of textsetting for each noun, but the data in 4.1. lack one due to the limitation mentioned above. The figures are arranged in order of the release date of the song in question with the oldest first. The bold vertical lines indicate bar lines in music score.

4.1. HL pattern: *sekai* ‘the world’

The figures below represent nine examples of the way the noun in question is set to musical beat in Japanese vocal music.

Figure 2: The grid notation of ‘sekai’ (lyric from *Sekaiha futarino tameni* sung by Naomi Sagara, 1967)

half notes	S:	x		x	
quarter notes	M:	x	x	x	x
eighth notes	W:	x x	x x x	x x x	x x
		se-		ka- i	

Figure 3: The grid notation of ‘sekai’ (lyric from *Sekaino kunikara konnichiwa* sung by Haruo Minami, 1967)

half notes	S:	x			x			x
quarter notes	M:	x		x	x		x	x
eighth notes	W:	x	x	x	x	x	x	x
		se-	ka-	i	(no)			

Figure 4: The grid notation of ‘sekai’ (lyric from *Blue light Yokohama* sung by Ayumi Ishida, 1968)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	
eighth notes	W:	x	x	x	x	x	x
			se-	ka-		i	

Figure 5: The grid notation of ‘sekai’ (lyric from *Tooi sekaini* sung by Akai Fuusen, 1968)

half notes	S:	x				x			x
quarter notes	M:	x		x		x		x	x
eighth notes	W:	x	x	x	x	x	x	x	x
		se-	ka-	i					

Figure 6: The grid notation of ‘sekai’ (lyric from *Ride on time* sung by Tatsuro Yamashita, 1980)

half notes	S:		x		x						
quarter notes	M:	x	x	x	x	x	x	x	x	x	
eighth notes	W:	x	x	x	x	x	x	x	x	x	x
								se-	ka-	i	

Figure 7: The grid notation of ‘sekai’ (lyric from *Sekaini hitotsudakeno hana* sung by SMAP, 2003)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:		x	x	x	x	x

Figure 8: The grid notation of ‘sekai’ (lyric from *Ano taiyoga, konosekaiwo terashi tsuzukeruyouni* sung by Kobukuro, 2011)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

Figure 9: The grid notation of ‘sekai’ (lyric from *Ashita sekai ga owarunara* sung by Mika Nakajima, 2012)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

Figure 10: The grid notation of ‘sekai’ (lyric from *Sugaono mamade* sung by Yuzu, 2014)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:		x	x	x	x	x

4.2. LH pattern: *sakura* ‘cherry tree/ blossoms’

Here follow ten examples of the way the noun in question is set to musical beat in Japanese vocal music.

Figure 11: The grid notation of ‘sakura’ (lyric from a traditional song, *Sakura sakura*)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

Figure 12: The grid notation of ‘sakura’ (lyric from *Sakurazaka* sung by Masaharu Fukuyama, 2000)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

Figure 13: The grid notation of ‘sakura’ (lyric from *Sakura (solo)* sung by Naotaro Moriyama, 2003)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:		x	x	x	x	x

Figure 14: The grid notation of ‘sakura’ (lyric from *Sakura (solo)* sung by Naotaro Moriyama, 2003)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:	x	x	x	x	x	x

Figure 15: The grid notation of ‘sakura’ (lyric from *Sakura* sung by Naotaro Moriyama, 2003)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

Figure 16: The grid notation of ‘sakura’ (lyric from *Sakura* sung by Kyogo Kawaguchi, 2003)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

Figure 17: The grid notation of ‘sakura’ (lyric from *Sangatsu Kokonoka* sung by Remioromen, 2004)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:	x	x	x	x	x	x

Figure 18: The grid notation of ‘sakura’ (lyric from *Sakura* sung by Ketsumeishi, 2005)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:	x	x	x	x	x	x

Figure 19: The grid notation of ‘sakura’ (lyric from *Sakura* sung by Ikimonogakari, 2006)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:	x	x	x	x	x	x

Figure 20: The grid notation of ‘sakura’ (lyric from *Sakura* sung by Ikimonogakari, 2006)

half notes	S:	x		x			
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
sixteenth notes	EW:	x	x	x	x	x	x

5. DISCUSSION

First, let us examine the case where Japanese HL accent patterns are set to musical beat. If the predicted correspondence between English ‘Strong-Weak’ and Japanese ‘High-Low’ worked in terms of textsetting, then a ‘High’ (*se-*) should be associated with a stronger beat than ‘Low’ (*ka-*). Three examples shown in Figures 2, 4, and 7 do present

this kind of correspondence, while the five examples shown in Figures 3, 5, 6, 8 and 9 do not. That is, 'High's (*se-*) are associated with weaker beats than 'Low's (*ka-*). At a first glance the last example shown in Figure 10 seems to support the correspondence since *se-* ('High') is associated with a stronger beat than *ka-* ('Low'). But *-i*, another 'Low', is associated with a stronger beat than *se-* ('High'). Therefore, the correspondence did not work. Secondly, let us examine the case where Japanese LH accent patterns are set to musical beat. Only the two examples shown in Figures 13 and 17 present the hypothesized correspondence. That is, 'High's (*ku-*) are associated with stronger beats than 'Low's (*sa-*). The remaining eight shown in Figures 11, 12, 14, 15, 16, 18, 19 and 20 are examples in which 'High's (*ku-*) are associated with weaker beats than 'Low's (*sa-*). Again, therefore, the correspondence is not supported. Accordingly, the apparent correspondence of English stress and Japanese pitch works only in one direction; while English stress may be equivalent to Japanese pitch height in production, pitch height is not equivalent to English stress in terms of textsetting.

In this connection, it will be of interest to examine how non-native speakers of English set text to tunes when they compose a song in English. Some ‘English’ songs composed using a VOCALOID™ (vocal + oid ‘like’) suggest a different textsetting from what the SDA predicts. Let’s take an example of one phrase *the world of coming together*, from a demo song [12]. Figure 21 is a scansion by the SDA, while Figure 22 shows the actual alignment translated into a grid used here.

Figure 21: Scansion by the SDA

half notes	S:		x		x		x				
quarter notes	M:	x	x	x	x	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x	x	x	x	x
		the	world	of	com-	ing	to-	ge-	ther		

Figure 22: Actual alignment found in a demo song

half notes	S:		x		x		
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x

the world of

(continued from the above grid)

half notes	S:	x		x	x		
quarter notes	M:	x	x	x	x	x	x
eighth notes	W:	x	x	x	x	x	x
			/ \		/ \		/ \
		co-	ming	to-	ge-	ther	

Part of the actual alignment, especially *world* and *toge-*, indicate a mismatch between stress and musical beat. Artists, of course, do not always follow the ‘rules’, but the stress-to-beat constraint is so strong in English that songs which ignore it would be difficult to follow.

6. CONCLUSION

This study examined whether a surface similarity between English stress and Japanese high pitch would be retained in the domain of textsetting. The examination of vocal music showed that the apparent similarity does not hold there. It was shown that Japanese HL pattern does not necessarily behave like English trochee, nor does Japanese LH pattern behave like English iamb in textsetting. Different ways of associating syllables with musical beat will show up when native speakers of different language background are faced with the task of textsetting.

7. ACKNOWLEDGEMENTS

This work is supported by a Grant-in-Aid for Scientific Research (C) (2) (Japan Society for the Promotion of Science, Grant No. 24520540).

8. REFERENCES

- [1] Dell, F., Halle, J. 2009. Comparing musical textsetting in French and in English songs. In: Aroui, J.-L., and Arleo, A. (eds), *Towards a Typology of Poetic Forms: from Language to Metrics and beyond*. Amsterdam/Philadelphia: John Benjamins. 63-78.
 - [2] Hayes, B. 2009. Textsetting as constraint conflict. In: Aroui, J.-L., and Arleo, A. (eds), *Towards a Typology of Poetic Forms: from Language to Metrics and beyond*. Amsterdam/Philadelphia: John Benjamins. 43-61.
 - [3] Lerdahl, F., Jackendoff, R. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
 - [4] Liberman, M. 1975. *The Intonational System of English*. The Indiana University Linguistics Club, Bloomington.
 - [5] Low, E., Grabe, E., Nolan, F. 2000. Quantitative characterization of speech rhythm: syllable-timing in Singapore English. *Language and Speech* 43, 377-401.
 - [6] McCawley, J. D. 1977. Accent in Japanese. In: Hyman, L. (ed), *Studies in Stress and Accent*, 261-302.
 - [7] Patel, A. 2008. *Music, Language, and the Brain*. Oxford/New York: Oxford University Press.
 - [8] Patel, A., Daniele, J. R., 2003. An empirical comparison of rhythm in language and music. *Cognition* 87, B35-B45.
 - [9] Patel, A., D. Iverson, J. R., Rosenberg, J. C. 2006. Comparing the rhythm and melody of speech and music: The case of British English and French. *J. Acoust. Soc. Am.* 119, 3034-3047.
 - [10] Taylor, E. 1989. *The AB Guide to Music Theory*. Part 1. London: The Associated Board of the Royal Schools of Music.
 - [11] <http://www.at-elise.com/> [Accessed 15 August 2014]
 - [12] http://www.crypton.co.jp/mp/pages/prod/vocaloid/mikuv3_bundle.jsp [Accessed 5 January 2015]