

Two Japanese Children's Songs

1. Purpose

This brief paper is an investigation of how the metrics of sung/chanted verse would work in a language that has no stress (Beckman 1986). I give background on the relevant aspects of Japanese phonology, then a structural description of two songs, together with their implications for metrical theory.

This paper is almost entirely a restatement of the content of an excellent senior honors essay written under my direction by Tami Swiger, a UCLA undergraduate; I've written the material from a viewpoint intended to express some of my current interests in the theory of meter. I have also included a bit of further data, a recording made for me by a native speaker of Japanese.¹

2. Background on Japanese phonology

2.1 *The Mora*

As a massive research literature indicates (see, e.g., McCawley 1968, Vance 1982, et seq.), Japanese prosodic phonology is very much centered on the *mora*. This is true for Japanese metrics as well. In the songs described here, the linguistic unit that mostly closely corresponds to the musical note is not the syllable, but the mora.

Here is a definition of the mora, with examples:

(4) A mora is any of the following:

a. Any short vowel, which may be preceded by a consonant:

Ueyama 'last name' consists of the moras *u e ya ma*.

b. Both "halves" of a long vowel: *Tookyoo* = *to o kyo o* 'Tokyo'

"Tokyo" in Japanese has two long vowels, which are sometimes Romanized, as here, with double letters. This is not just spelling; a Japanese long vowel really is pronounced twice as long as a short vowel. Note: "ky", "ry" count as just one sound in Japanese.

c. A nasal sound at the end of syllable: *honda* = *ho n da* 'Honda'

d. The first half of any long (double) consonant: *Tottori* = *to t to ri* 'name of city in Japan'

¹ I would like to thank Motoko Ueyama, at the time a graduate student in the Applied Linguistics/TESL Department at UCLA, who kindly (and quite ably) sang the two songs on tape.

The words above could also be divided into syllables: *u e ya ma, too kyoo, hon da, Tot to ri*. Thus, many syllables include two moras (and few even contain three).

At least as far as I can tell from my own elicitation, Japanese speakers “feel” a rhythmic beat on every mora of a word. Thus *Tookyoo* is felt to have four rhythmic beats (*to o kyo o*), whereas an English speaker listening to the very same word (in correct Japanese pronunciation) hears just two beats: *too kyoo*. The same holds true for moras that are acoustically very quiet or even silent: a Japanese speaker will hear *Tottori* as *to t to ri*, with four beats. This has consequences for music. Because Japanese speakers are more closely attuned to moras than syllables, they tend to feel that every mora counts as (at least one) note of the music, just as English speakers feel that every syllable counts as at least one note.

To clarify: for English, sometimes one syllable can have more than one note. This happens if the musical pitch changes during the syllable (try for example the syllable *ev* of *ev'ry* in “My Country 'Tis of Thee”). But if you do the reverse, and sing two syllables on the same pitch, (for example, syllables 1-2, 3-4, 5-6 of “Twinkle, Twinkle, Little Star”) it sounds like two notes (and is written as such in standard musical notation). Japanese is just the same, but moras instead of syllables. If two moras are sung consecutively on the same pitch, Japanese speakers will always hear it as two notes (and it will be written down as such by Japanese-speaking musicians). This mode of perception is found even if the two moras belong to the same syllable, as in the two syllables of *Tookyoo*.

In what follows, I will write down the Japanese words mora by mora, to depict this system. I will also discuss some perceptual illusions to which non-Japanese speakers are vulnerable when they listen to Japanese songs.

2.2 Absence of Stress

Japanese has no stress. It is pronounced in a steady, even rhythm, without giving any special force to any particular syllable or mora (Beckman 1986).

This poses an important question: if there is no stress to mark off the rhythmically strong syllables, how can you have rhythm in songs? In a stress language like English, the stresses are crucial in signaling the rhythmic structure. We will see a possible answer to this problem later on.

Japanese does have something that is vaguely like stress, namely its so-called “pitch accent”. This means that in a particular word, one of the syllables has a pitch that is noticeably higher than the pitch of the other syllables. Thus, *sake*, with the first syllable higher in pitch, means “salmon”; *sake*, with second syllable higher in pitch, means “rice wine”.

As far as I can tell, these pitch differences are apparently entirely irrelevant to singing: there is virtually no tendency for the high-pitched syllable to act as if it were stressed and thus occur on strong beats. I will therefore ignore pitch accent in what follows.

3. "Tooryanse"

This song is the accompaniment to a children's game, in which two children form a "gate" by holding hands with arms lifted, and other children pass through the gate. It describes a woman conversing with a gate-keeper, who is barring the way to a temple of the god Tenjin, to which the woman wishes to go to offer a talisman (religious token) in celebration of her child's seventh birthday.

A tiny bit of cultural background: fiercely guarded gates apparently play an important role in Japanese history. Children's seventh birthdays are considered especially significant in Japan.

3.1 Transcription of phonological juncture

The transcription is given with a rough estimate of the degree of phonological juncture (Selkirk 1980, Hayes 1989). The symbols are as follows.

- (5) ### right edge of an Intonational Phrase
 ## right edge of a Phonological Phrase
 # right edge of a Clitic Group
 = right edge of a Word

3.2 Meter

I assume that the meter consists of a metrical grid (Lieberman 1975, Lerdahl and Jackendoff 1983). This grid is dipodic, in the sense of Prince 1989) and consists of four dipods, each dipod being trochaic on both upper and lower levels. The grid material for a single dipod is shown below.

```

x
x      x
x    x  x    x

```

It is assumed that the dipods have an internal structure (two binary feet), and are grouped into a higher level structure (Attridge 1982, Hayes 1988, Hayes and MacEachern 1998): two dipods form a hemistich (half line) and two half-lines form a line.

```

x      x      x      x
x    x  x    x  x    x  x    x
x x  x x  x x  x x  x x  x x  x x
[  ] [  ] [  ] [  ] [  ] [  ] [  ] [  ] feet
[      ] [      ] [      ] [      ] dipods
[                ] [                ] hemistichs
[                                ] line

```

From Tami Swiger's paper, it appears that very many Japanese children's songs employ this grid.

In some cases, the lines are further grouped into couplets.

3.3 Text

Here are the words of the song, with attempted word-by-word and idiomatic translations.

(6)

- 1- Tooryanse ### tooryanse ###
one-may-pass-through one-may-pass-through
- 2- koko =wa ## dokono ## hosomichijya ###
here TOPIC which small-path
"What is this small path?"
- 3- tenjinsamano # hosomichijya ###
toward-Tenjin small-path
"It is the path to (the god) Tenjin."
- 4- chootto # tooshite ## kudashyanse ###
a-little pass please
"Please let me pass through."
- 5- goyoononaimono ## tooshasenu ###
business-POSSESSIVE-not-exist-person pass-not-allowed
"A person who has no business here may not pass."
- 6- kono # ko =no ## nanatsu =no ## oiwai =ni ###
this child POSS 7-years-old POSS celebration DATIVE
"For the seventh birthday celebration of this child,"
- 7- ofuda =o ## osameni ## mairimasu ###
talisman OBJECTIVE contribute I-am-going
"I am going to contribute a talisman (religious token)."
- 8- iki =wa ## yoi # yoi ### kaeri =wa ## kowai ###
going TOPIC good good return TOPIC frightening
"The going is pleasant; the return is frightening."
- 9- kowainagaramo ## tooryanse ###
frighten-although one-may-pass-through
"Although it is frightening, one may pass through,"
- 10- tooryanse ### one-may-pass-through

3.4 Scansion

Next, I have provided an alignment of the moras (not syllables) to a rhythmic grid; taken from Swiger and checked against my recording. Below the grid are brackets indicating quarter-lines, hemistichs, and lines. There were obtained by counting off grid marks at intervals of 4, 8, and 16.

(7)

-1- x x x x
 x x x x x x x x
 x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 To o rya n se ### to o rya n se ###
 [] [] [] []
 [] [] []
 [] []

-2- x x x x
 x x x x x x x x
 x x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 ko ko=wa ## do ko no ## ho so mi chi jya ###
 [] [] [] []
 [] [] []
 [] []

C
O
U
P
L
E
T

1

-3- x x x x
 x x x x x x x x x
 x x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 ten ji n sa ma no # ho so mi chi jya ###
 [] [] [] []
 [] [] []
 [] []

-4- x x x x
 x x x x x x x x
 x x x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 cho o t to # to o shi te ## ku da shya n se ###
 [] [] [] []
 [] [] []
 [] []

C
O
U
P
L
E
T

2

-5- x x x x
 x x x x x x x x
 x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 goyo o no na i mo no ## to o shya se nu ###
 [] [] [] []
 [] [] []
 [] []

-6- x x x x x
 x x x x x x x x
 x x x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 kono # ko =no ## na na tsu =no ## o i wa i =ni ###
 [] [] [] []
 [] [] []
 [] []

C
O
U
P
L
E
T

3

-7- x x x x

```

x      x      x      x      x      x      x      x      T
x x x x  x x x x  x x x x  x x x x
| | | |  | | | |  | | | |  |
o fu da=o ## o sa me ni ## ma i ri ma  su ###
[          ][          ][          ]
[          ][          ]
[          ]

```

```

-8-  x      x      x      x      x      x      x      x
x x x x  x x x x  x x x x  x x x x
| | | |  | | | |  | | | |  | | | |
o i ki=wa ## yo i # yo i  ### ka e ri =wa ## ko wa i ###
[          ][          ][          ][          ]
[          ][          ][          ]
[          ]
T
R

-9-  x      x      x      x      x      x      x      x
x x x x  x x x x  x x x x  x x x x
| | | |  | | | |  | | | |  | | | |
o kO wa i  na ga ra mo ## to  o  rya n se ###
[          ][          ][          ][          ]
[          ][          ][          ]
[          ]
L
E

```

```

-10- x      x      x      x      x      x      x      x      T
x x x x  x x x x  x x x x  x x x x
| | | |  | | | |  | | | |  |
to o rya n se ###  o  o
[          ][          ][          ]
[          ][          ]
[          ]

```

3.5 Hierarchical structure

All the lines in the song fill all four of the strongest positions in the grid; in the terminology of Burling (1966) and Hayes and MacEachern (2000) they are “4-beat” lines. The exception is the last line, a curious “2”.

There seems to be an overall structural organization above the level of the line, that seems to go along with the structure of the dialogue depicted in the song. Roughly, there is an introductory line (line 1), three similar couplets (2,3, 4,5, 6,7), and a concluding triplet (8-9-10).

The opening line, *Tooryanse, tooryanse* differs from what follows in that it has a couple of “slow moras”; that is, moras set to two instead of one grid position. It “sets the scene” for the dialogue that follows.

After the opening line, there are three dialogue couplets. The first two are single exchanges, and the third is a single speech by the mother. The couplets all have nearly the same arrangement of moras, in which every or almost every position in the grid is filled except for the last three, which are left empty. Moreover, the couplets have the same contour of pitches. The

pitch contour is set at a lower overall level in the second of the three couplets (in musical terms, it is transposed down a fourth), then in the third couplet it goes back to its original pitch.

The concluding triplet consists of two very similar lines (which might be thought of as a couplet within the triplet), followed by a very short 2-beat line. The two similar lines introduce a new, tenser mora rhythm: the very first position in the line is left empty, but the moras now extend all the way to the fifteenth position in the line.

There is a slowing of the mora rhythm (that is, unfilled grid positions) at the first of the two final repetitions of *tooryanse*. These repetitions tend (at least as I intuit it) to integrate the final *tooryanse* with the preceding two lines, forming a complete triplet. Note that the word *tooryanse* acts as a kind of “frame” for the song, introducing and concluding the action.

There's not too much to say about how this formal structure relates to the laws discussed in class. The couplets are all uncontroversial 4-4's, obeying the Couplet Law. The final triplet is 442, which certainly seems rather like a 443 triplet, which would obey a generalized version of the Couplet Law. Lastly, if one accepts the view that the last three lines form a triplet, it fits in with the general esthetic principle that long elements should come late (triplet follows couplets).

3.6 Perceptual Illusions

The patterning of beats in the grid illustrates the difference between syllables and moras, along with the differing phonological perceptions of English and Japanese speakers. Since the English speaker is liable to hear syllables instead of moras, (s)he will tend to hear some rhythmic configurations that probably weren't intended in the Japanese original.

For example, I personally tend to hear the sequence *goyoono* (beginning of second line of Couplet 2) as having a syncopated rhythm; a sort of Early Medium, as in (7). This is apparently a perceptual illusion: if one listens correctly, attending to moras instead of syllables, the sequence is not syncopated at all, but simply involves one linguistic unit per beat, as shown in (8):

| | | | |
|-----|--|-----|---|
| (7) | x x x x x x x go yoo no | (8) | x x x x x x x go yo o no |
|-----|--|-----|---|

3.7 Syllable Durations

Virtually every mora gets exactly one beat. There are a few exceptions. Line-final syllables (which always precede ###) are sung long. This makes sense; since they precede a strong break, these syllables would be likely to be long in normal pronunciation as well.

There are two other places in which a mora gets one beat: (a) the cases already mentioned, in which there is slowing on the first two moras of *tooryanse*, probably as a framing device for the overall structure of the song; (b) The two Strong-is-Longs mentioned above (*koko=wa* and

dokono) are probably long to satisfy a requirement of fitting phonological breaks to metrical breaks; this is discussed below.

An interesting consequence of cases like *koko=wa* and *dokono* is that you can't learn the proper vowel lengths in Japanese words just by listening to songs; you have to hear them in normal speech. The reason is this: if in a song you hear a vowel that lasts for two beats, it might last two beats because it is a genuine long vowel (in spoken Japanese), but it also might be two beats long because the song happens to give two beats to a single mora.

3.8 Matchup of phonological breaks with metrical breaks

In the chart below, I've indicated the distribution of phonological breaks according to where they occur in the structure of the line (that is, at line breaks, at hemistich breaks, at quarter-line breaks, and otherwise). I first give the raw numbers, then convert them by rows to a percentage for clarity.

| | | | | | |
|-----------|------------|-----------|-----------|----------|-----------|
| (9) | ### | ## | # | = | none |
| Line | 10 | 0 | 0 | 0 | 0 |
| Hemistich | 2 | 6 | 3 | 0 | 0 |
| Dipod | 0 | 5 | 1 | 1 | 11 |
| Other | 0 | 0 | 1 | 6 | 72 |
| %: | | | | | |
| Line | 100 | 0 | 0 | 0 | 0 |
| Hemistich | 18 | 55 | 27 | 0 | 0 |
| Dipod | 0 | 28 | 6 | 6 | 61 |
| Other | 0 | 0 | 1 | 8 | 91 |
| | ### | ## | # | = | none |
| Line | 10 | 0 | 0 | 0 | 0 |
| Hemistich | 2 | 6 | 3 | 0 | 0 |
| Dipod | 0 | 5 | 1 | 1 | 11 |
| Other | 0 | 0 | 1 | 6 | 72 |

I would say that the agreement is remarkably good. There are no cases in the “bad” corners of the chart (these would have maximal disagreement of phonological and metrical breaks), nor are there even any that are near the bad corners. Across line boundaries, we always get the largest possible phonological break; and within quarter-lines (the tightest metrical domains), we get absence of break (i.e. moras within the same word) over 90% of the time. The intermediate cases (moras that occur across a hemistich or quarter-line break) have suitable intermediate values.

3.8.1 Statistical testing

A simple way to determine whether the disparities seen in the chart could be the result of chance is to divide it up in all possible 2 x 2 arrangements, count the aggregated numbers in each of the four resulting cells, and perform a chi-square test on the result. For instance, here are three possible divisions:

| | ### | ## | # | = | none |
|-----------|-----|----|---|---|------|
| Line | 10 | 0 | 0 | 0 | 0 |
| Hemistich | 2 | 6 | 3 | 0 | 0 |
| Dipod | 0 | 5 | 1 | 1 | 11 |
| Other | 0 | 0 | 1 | 6 | 72 |

yielding: 10 0
 2 106

| | ### | ## | # | = | none |
|-----------|-----|----|---|---|------|
| Line | 10 | 0 | 0 | 0 | 0 |
| Hemistich | 2 | 6 | 3 | 0 | 0 |
| Dipod | 0 | 5 | 1 | 1 | 11 |
| Other | 0 | 0 | 1 | 6 | 72 |

yielding: 21 0
 7 90

| | ### | ## | # | = | none |
|-----------|-----|----|---|---|------|
| Line | 10 | 0 | 0 | 0 | 0 |
| Hemistich | 2 | 6 | 3 | 0 | 0 |
| Dipod | 0 | 5 | 1 | 1 | 11 |
| Other | 0 | 0 | 1 | 6 | 72 |

yielding: 21 0
 14 83

All twelve divisions, submitted to a 2 x 2 chi-square test, yield highly significant results, with p no higher than $4 \times 10^{-7^2}$ in any comparison.

3.8.2 Assessment

This is the best agreement of juncture and line structure that I have ever seen. It suggests an answer to the question asked earlier: what do you do to make a line sound rhythmic when your language doesn't have stress, and thus cannot directly manifest strong and weak beats? The answer seems to be that you make far more use of the principle of break match-up. A line of a song in Japanese sounds rhythmic to Japanese speakers because it is very carefully matched in its phonological breaks to the appropriate metrical breaks.

² Note: these need to be redone with the Yates correction, which will make them slightly higher.

3.9 Two-Mora Syllables

Although the basic assignment of linguistic material to grid positions seems mainly to work by moras, there does seem to be a limited role played by syllables as well.

Consider that when a word has two moras in a row, sometimes they will form separate syllables (for example *mo to* ‘proper name’) and sometimes will form just one syllable (for example, the first two syllables of *too ryan se*). The question is whether the grouping of moras into syllables has consequences for Japanese metrics.

Apparently it does. Suppose we take all the two-mora syllables in “Tooryanse” and examine in each case if the first mora is in a stronger grid position than the second, or in a weaker position. Here are the configurations I have in mind, for the sample syllable yan:

(10)a. First Mora in Stronger Position b. Second Mora in Stronger Position

| | | | | | | | | | | |
|----|---|----|----|---|----|----|---|----|---|---|
| x | | | | x | | | | | x | |
| x | | x | | x | x | | | | x | x |
| x | x | or | x | x | or | x | x | x | x | x |
| | | | | | | | | | | |
| ya | n | | ya | n | | ya | n | ya | n | n |

We find that cases where the first mora is placed in a stronger position are quite numerous: above, these are (in order) too, ryan, too, ryan, ten, jin, choo(t), too, shyan, nai, too, oi, wai, dao, mai, yoi, yoi, kae, wai, too, ryan, too, and ryan (23 total). There are only two cases in the whole song in which the second mora is in a stronger position than the first: yoo (line 5) and wai (line 8).

Clearly, this degree of disparity could only arise if it involves a genuine structural principle of the system. We might speculate that the first mora of a two-mora syllable somehow is felt to be phonologically “stronger” than the other, and thus more deserving of being placed in a strong metrical position. Such a principle makes quite a bit of sense for two-mora syllables like *yan* or *tot*, since a vowel is louder than a consonant. The principle makes less sense when the two moras have identical vowels, for example in the syllable *too*.

4. A Simpler Song, with a /sh/ Mora

A second song has yet to be fully analyzed. (I believe a full analysis would basically support the conclusions drawn above.) Two main points emerge from this song:

(a) The Japanese tradition does allow quatrains, though it is probably less fond of the quatrain than the English folk tradition is.

(b) Sometimes a mora can be surprisingly non-sonorous (acoustically weak). The boldface mora below is simply the first half of a long /ʃ/ sound. Nonetheless, in the Japanese system it counts as a legitimate beat, and is “sung” (even though /ʃ/ cannot support pitch) with the proper amount of time.

The musical scale used in this song is the pentatonic, used widely in the folk songs of the Southern Appalachians and elsewhere around the world.

(11)

- 1- Yuuyake ## koyakede ## hi =ga ## kurete ###
 evening-burning little-burning sun NOMINATIVE setting-and
 "As the sun sets,"
- 2- yama =no ## otera =no ## kane =ga ## naru ###
 mountain POSS temple POSS bell NOMINATIVE ring
 "The bell of the mountain temple is ringing."
- 3- otete ## tsunaide ### mina ### kaero ###
 little-hands holding everybody let's-go-home
 "As we are holding hands, let us go home."
- 4- karasu =to ## isshoni ### kaerimasho ###
 crow with together let-us-go-home
 "Together with the crows, let us go home."

(12)

- 1- x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 Yu u ya ke ## ko ya ke de ## hi=ga ## ku re te ###
- 2- x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 ya ma=no ## o te ra =no ## ka ne=ga ## na ru ###
- 3- x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 o te te ## tsu na i de ### mi na ### ka e ro ###
- 4- x x x x x x x x x x x x x x x x
 | | | | | | | | | | | | | | | |
 ka ra su=to ## i ſ ſo ni ### ka e ri ma sho ###

5. Conclusion

The salient aspects of the Japanese metrical system, at least as can be seen from limited data, are:

- The use of the mora, rather than the syllable, as the basic rhythmic unit.
- The absence of stress as basis for making the rhythmic structure apparent.

- An (apparently compensating) powerful tendency to align phonological breaks with metrical breaks.
- The use of a principle to the effect that the first mora of a two-mora syllable “counts as” stronger than the second mora, and should normally be placed in a stronger grid position.

References

- Attridge, Derek (1982) *The Rhythms of English Poetry*. xxx
- Beckman, Mary E. (1986) *Stress and Non-Stress Accent*. Walter de Gruyter
- Burling, Robbins (1966) “The metrics of childrens’ verse: a cross-linguistic study” *American Anthropologist* xxx
- Hayes, Bruce (1988) “Metrics and phonological theory,” in Frederick Newmeyer, ed., *Linguistics: The Cambridge Survey*, Cambridge University Press, Cambridge, 220-249.
- Hayes, Bruce and Margaret MacEachern. 1998. Quatrain form in English folk verse. *Language* 74:473-507.
- Hayes, Bruce. 1989. The Prosodic Hierarchy in meter. In Kiparsky and Youmans (1989), pp. 201-260.
- Lerdahl, Fred and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- Lieberman, Mark (1975) The intonational system of English. Dissertation, MIT. On line at MIT.
- McCawley, James (1968) *The phonological component of a grammar of Japanese*. Mouton.
- Prince, Alan S. (1989) “Metrical forms,” in Kiparsky and Youmans (1989).
- Selkirk, Elisabeth O. (1980) “Prosodic domains in phonology: Sanskrit revisited,” in Mark Aronoff and Mary-Louise Kean, eds., *Juncture*, Anma Libri, Saratoga, CA.
- Vance, Timothy (1982) *The phonology of Japanese*. SUNY Press. xxx