The relationship between note values and speech timing ― Moraic representation in Japanese children's songs

Hitomi Nakata

School of Linguistics & Applied Language Studies, The University of Reading

Abstract. This study examines the temporal relationship between a language (Japanese) and the music written by and for speakers of that language. It asks whether that musical timing reflects language timing. In order to investigate this question, we look here at the relationship between the temporal value of musical notes and the mora structure of words which are represented as lyrics in Japanese children's songs. After analyzing and discussing some previous studies, I will present data from 60 songs (*warabeuta* – traditional children's songs and *dooyoo/shooka* – less-traditional children's songs), focusing on three non-standard moras (hereafter, NS moras: the first member of geminates: Q; a lengthened vowel: R; and a moraic nasal: N) which often behave as a part of bimoraic unit. The results suggest the following: 1) a one-mora unit (CV) tends to be mapped onto shorter notes in modern songs which are composed using a non-pentatonic (Western) scale; 2) a two-mora unit (CVμ), especially CVQ and CVN exhibit a strong tendency to be realized on a single note, which leads to the assumption that they are monosyllabic, contrary to the traditional framework of Japanese sound structure; 3) the influence of loan words in modern music can be issues for future studies. The results of the present study, therefore, will be the basis of experiments aimed at providing evidence for possible principles of rhythm in speech and music.

1. Introduction

It can be hypothesised that there is a temporal relationship between a language and the music written by and for speakers of that language, i.e. that musical rhythm should reflect language rhythm. In order to investigate this question, we look here at the relationship between the temporal value of musical notes and the mora structure of words which are represented as lyrics in songs. In this paper, we will observe some previous studies in which this relationship is discussed and analysed. After examining and discussing the studies, then, I will present a new set of data for which the interaction of lyrics and musical notes are investigated.

Reading Working Papers in Linguistics 8. 69-93.

1.1 Background-1: Empirical study in English

The issue of the relationship between music and speech has been widely discussed (e.g. Patel & Daniel 2003, Risset 1991, Selkirk 1984), and the majority of studies on the rhythmic relationship between music and speech have taken only one approach, i.e., analysing existing songs. One of a few experimental studies is by Palmer & Kelly (1992). They examine how metrical rules in English (e.g. iambic: weak-strong-weak) and metrical rules in music (i.e. the first and the third beats are accented) are notated in music. The correspondence between linguistic and musical accentuation in their study was investigated from the perspectives of prosodic structures beyond the unit of syllables and/or words, under some grammatical constraints. The study is based on both analyses of composed/notated songs and experiments in song performance. Palmer & Kelly in fact carried out two sets of studies: each set consists of analysis of both song composition and performance of vocalists in experimental settings. The four studies are summarised as follows; 1) analysis of English song composition: investigation of whether Compound and Nuclear Stress rules[1] are represented in musical meter, 2) test of vocalists' performance of 1), 3) analysis of composition: if violation of 1) (e.g. a phrase having a clitic accent such as 'the black bird' (vs. 'black birds') changes the stress, and 4) test of vocalists' performance of 3).

Table 1 shows the examples of results from the first experiment. (The title and names of the items have been adjusted from Palmer & Kelly 1992:528).

Table 1 The relation between linguistic and musical metrical accent

	Musical accent	
Linguistic accent	1st beat	2nd beat
In Gilbert & Sullivan operettas (words and music set together)		
1st syllable (compound words)	43	5
2nd syllable (2-word phrases)	16	79

[1] In the compound word 'blackbird', assignment of primary stress is placed upon the first [a], whereas in the phrasal words 'black bird', the primary stress is assigned to [i] of 'bird' (Chomsky & Halle 1968).

Results showed a high rate of assigning musical accents to the first syllable of compound words and to the second syllable in two-word phrases. Further experiments with vocalists in the second study demonstrated similar results, i.e. singers tend to place longer durations on syllables where both linguistic and musical accents coincide. The third and the fourth studies exhibited similar results; e.g. “lake” in a clitic phrase “the cold lake” tends to be aligned (and sung) with a strong beat in music although the tendency is weaker than that of “cold lakes” due to the existence of the clitic “the”.

This work is noteworthy in that it is evidence of linguistic accentuation providing a strong influence on syllable duration over certain grammatical conditions when represented as song texts within a musical framework. These studies are unusual in that they deal with the relationship between language and music using both analysis of composed songs and experiments and which are objective rather than impressionistic.

However, several questions can be raised. First, they measure only durations to compare accented syllables and accented notes in music, but stress involves intensity which equally deserves experimental investigation. Second, this study does not take account of qualitative/durational differences among syllable types in English. For example, is there a tendency to link a “short” CV syllable with a quaver or crotchet? Is a syllable with more phonological units such as a CVCCC syllable linked to longer notes? What about the quantitative difference between vowels (short vs. long vs. diphthongal)? These questions investigate the phonetic and phonological rather than the grammatical aspect of temporal correlations between linguistic units and musical notes.

1.2 Background-2: Non-empirical/empirical studies in Japanese

The composer Yamada (1886~1965) pioneered the study of rhythmic interrelations between Japanese speech and music. According to Dan (1999), who is also a composer, Yamada first developed a theoretical approach to song texts with consideration of the accentuation system in Japanese. Yamada also advocated a principle of “one (speech) sound unit per note” (in which one (speech) sound unit can be assumed as one mora). This approach is considered as tradition in composing Japanese songs and was investigated systematically by Yamada.

Vance (1987) also supports the “one-mora per beat” principle in Japanese music. He cites the case of traditional songs which consist of only CV or V in the lyrics. For example, *Urashima Taroo* (a boy’s name in a fairy tale) – an anonymous song – has 47 moras for the first verse, all moras are either CV or V, and all are assigned one beat without exception. Vance mentions several other examples of traditional children's music in which phrases maintain one-note/one-mora relations.

Sakai (1998) gives duration measurements of geminated consonants in song texts. She focuses on the durational properties of geminated unvoiced consonant stop (Q) in speech and in singing, and finds that the ratio of the preceding vowel and the geminated stop is about 7:3 both in speech (of children) and singing of children's songs: *warabeuta* and *shooka.* She emphasises that this ratio is in line with studies on geminated consonants of adults in speech, concluding that an unvoiced stop is recognised as the first part of a geminated consonant *if* it occupies more than 30% of the duration of the preceding vowel.

Sakai (1998) thinks that analyses solely from written music are problematic, because the style of written materials are varied according to composers, age of composing, tempo, type of music, and purpose of publication, etc. However, the performance will also vary time to time for any given singer or across singers. In other words, live performance can vary unpredictably and need not reflect a fixed underlying pattern. Even though the performance sounds perfect, the criteria for perfection in performing music is unclear, and perhaps the perfect performance simply does not exist. We can only observe the outcome of execution of written texts which can be stylised. But this can be avoided if songs are collected from a standardised compilation of music.

2. Tanaka's study

Tanaka (2000) attempts to investigate the relationship between the mora, the syllable, and the duration of musical notes including more modern songs (created after 1960). He reports that the one-to-one relationship noted by Vance (1987) is not found so regularly in NS-moras which are often attached to the prior standard mora (hereafter, S-mora). He hypothesises that the S+NS mora forms a unit of rhythm and behaves like a heavy syllable as a speech unit, based on his findings of the low rate of assigning NS-moras to one note (Q:52%; R:30%; and N:22%). He further speculates that the assignment of a mora on a musical note reflects the rhythm patterning of a language, and hypotheses that the concept of the syllable as a unit of speech segmentation plays more important role in the (modern) standard Japanese.

2.1 Material

Among studies of Japanese song-data, Tanaka's (2000) must be one of the most extensive. From 866 songs, he has isolated 7461 words and 21335 moras. His 866 songs are from a compilation called *Purofessyonaru yuusu: Kayookyoku no subete besuto 1100* (*Professional use: best songs of all pop music 1100*, Zen'on Publishers 1999). He first categorises these into two groups' *enka* and *hi-enka*. *Enka* in Japan is a genre of traditional pop songs and *enka* singers usually use a lot of tremolo to express certain emotions. *Hi-enka* ('non-*enka*') therefore means songs which do not fit into that categorisation and are usually similar to Western pop music. Secondly, Tanaka divided all the target songs into seven groups according to in what year each song was composed, ranging from 1930 to 1998, as seen below in Table 1. Thirdly, only nouns are employed, and sorted into three groups: native Japanese, Sino-Japanese, and loan words.[2] Finally, the moras were sorted into two groups: standard (CV and V) and non-standard (Q, R, and N) moras (S-moras and NS-moras, respectively). The following table shows the compiled results of all these classifications. (The title and the items in columns are translated from Tanaka 2000:154).

Table 2 The number of moras, song notes, and the ratio of a mora to a note in Japanese songs of each decade (1930-1998)

Period of composition	No. of songs	No. of nouns	No. of moras	No. of notes	Notes per mora	No. of notes per S-mora	No. of notes per NS-mora
1930 – 39	73	462	1330	1700	**1.28**	**1.29**	**1.13**
1940 – 49	74	588	1641	1887	**1.15**	**1.56**	**1.07**
1950 – 59	89	671	1933	2186	**1.13**	**1.15**	**0.99**
1960 – 69	180	1281	3130	3399	**1.09**	**1.10**	**0.96**
1970 – 79	185	1698	5000	5129	**1.03**	**1.04**	**0.90**
1980 – 89	149	1533	4526	4660	**1.03**	**1.06**	**0.79**
1990 – 98	116	1228	3775	3806	**1.01**	**1.04**	**0.69**
Total	866	7461	21335	22767	**1.07**	**1.09**	**0.88**

[2] He does not mention it clearly but the majority of loan words are assumed to be English.

2.2 Data analysis: rhythmic shift in mora value per one note

Tanaka observes that the number of musical note per mora decreases as years go by for both S-moras and NS-moras as seen in the three columns in bold. In other words, one note bears more moras over the years, i.e. more speech is compressed into a unit of music in later songs. One exception is the cross section of “Number of notes per S-mora” and “1940-1949” (which is shaded and shows one CV bears 1.56 moras per note). It seems that the value of S-moras to musical notes increased during the war years. It is beyond the scope of this study to investigate what sort of changes occurred in Japanese songs in that period, but it was then that almost all English songs were banned and military songs were encouraged by the authorities up to 1945 (Dan 1999). This sort of social control might have had an effect, and it is certainly the case that the situation reversed itself after the war.

Tanaka’s account of rhythmic changes suggests that, at one time, the mora could be sung over two notes of different pitches, but gradually in later years the one-mora/one-note relationship became more usual. He further observes that the one-to-one relationship between moras and notes is not found in newer songs, i.e., that one note sometimes maps into two moras or more. These changes are sketched in the following figures (1)-a. and (1)-b., which are restructured from Tanaka (2000:155).

(1) Factors of rhythmic changes

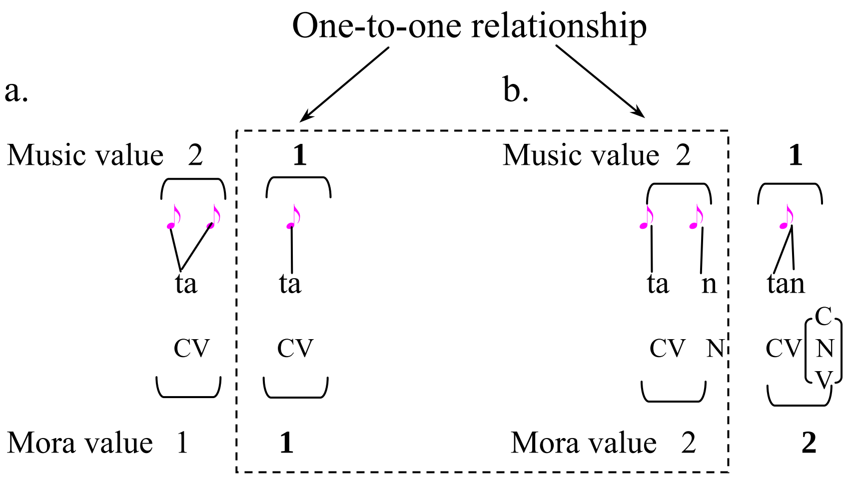
a. Decrease of melody-shift (to one-to-one relationship)

b. Change of the relationship (to more moras per note)

Here, Tanaka uses ♪ (quaver) as a symbol of any note that is independent of time; i.e., he does not specify how long a note can be, but uses the symbol as a unit of one note. These two changes of assigning moras to notes can be elaborated and exemplified below with some example moras ([ta] for CV and [n] for N), and the relative value (♪= 1) for each case.

[3] Here ‘μ’ represents any mora including a non-CV.

(2)

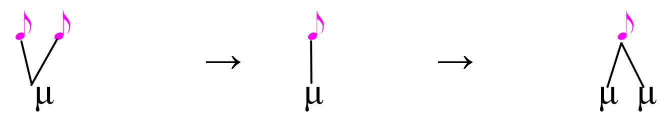


Note that in (2)-a, the mora value remains **as one** while the notes decrease from two to one, which leads to a one-to-one relationship between mora and note. In (2)-b, it starts from the same one-to-one relationship, and the number of notes shifts from two to one, but the number of moras remains **two**. The rightmost (2)-b. bears mora values as much as four times more than the leftmost pattern of (2)-a, moving from 1/2 in the old period to 2/1 in newer songs. However, it should be noted that in (2)-b, the second mora is NS, which may not count as the durational equivalent of an S-mora. It is questionable to directly compare the note value to the moraic value without considering the measured duration of each unit. Tanaka does not point this out in his study, but it is an essential part of the diachronic perspective.

2.3 Summary and limitation of the study

We have observed how moras behave in the context of musical manifestation in three aspects: historical period, type of mora, and type of music. Tanaka suggests that there has been a durational shift toward bearing more moras per note or fewer notes per mora in all these categories, despite some particular phenomenon in each sub-categorisation.

(3)



Period :	old songs	→	new songs
Type of moras:	S-moras	→	NS-moras
Type of words:	native Japanese	→	loan words
Type of music:	*enka* (traditional)	→	*non-enka* (modern)

Tanaka makes the very interesting observation that even CVCV combinations are assigned to a single mora in songs composed after the 80s. He concludes that the syllable or bisyllabic sequence (CVμ and/or CVCV) plays more of a unitary role because in modern songs they are assigned to one note.

Here two things can be reconsidered. First, Tanaka's definition of *enka* is not clear. He categorises *enka* songs as traditional hence non-Westernised Japanese songs, and the analysis was made on the dichotomy of *enka* and non-*enka* in which non-*enka* songs (modern with a touch of Western style) exhibit the phenomenon of assigning CVμ onto one note. However, he does not give any criteria of difference between *enka* and non-*enka*, from either musical structure (e.g. melody patterning or rhythmic features) nor linguistic factors in lyrics. In fact, Kindaichi (1967) classifies *enka* songs (although he terms the group in different way: *ryuukoo-kyoku* 'popular songs' and *shin-minyoo* 'new folk songs') as a Western song group, saying that '… this type of songs adopt an unrestricted style, using both Western and Japanese melody lines. About half of *enka* songs are Japanese style, and many employ minor modes such as the Ionian scale'(ibid.:439). According to this definition, the borderline of *enka* and non-*enka* seems to be ambiguous, hence it seems to be difficult and perhaps even meaningless to divide songs into these two groups.

Second and more importantly, the main concern of this study seems to be the number of moras and notes as units of beat, but not on the durational relationship. Although Tanaka briefly mentions in the end that '… a CVμ mora somewhat tends to be assigned to a longer note than a CV is …' (ibid.:35-36), it appears that the focus is placed on the notion of one musical unit as a bearer of a mora or moras. However, I assume the temporal value (both of speech and of music) is indispensable to investigate a rhythmic phrase; i.e. a quantitative relation should be the crucial factor. It is then of interest that how often a NS-mora is mapped on to a (possibly, longer) note, both independently and dependently (as μ of CVμ).

He suggests that one possible reason for the shift towards more moras per note is adopting more loan words from English, which allows consonant clusters, into modern pop songs. However, since he pays little attention to musical duration for each unit, his argument is ambiguous in terms of possible rhythmic interactions between speech and music. Durational concern in both speech unit and musical unit needs to be considered when temporal organization is examined, rather than discussing the correspondence of number of units in two domains.

3. Nakata's study on children's songs

To explore the areas mentioned above, I collected another set of data from children's songs. The musical timing value of notes was considered as well as the linguistic value of mora units when examining the assignment of a speech unit onto a musical unit, and an attempt was made to establish the statistical significance of the observed data. The quantity of data is more limited than Tanaka's, but attempts were made to investigate the durational relationship between notes and moras more closely in detail. While using a similar approach and methodology to Tanaka's, this study places emphasis on the quantitative differences among moras, notes and song groups.

In order to examine the relationship of rhythmic patterning, traditional children's songs (or *warabeuta* in Japanese) are particularly suitable. Some linguists and musicologists claim there is a close relationship between spoken words and lyrics represented in *warabeuta* songs. From the linguistic point of view, Kindaichi says '*Warabeuta* reflects the accent patterns in Japanese, i.e. melody in music often represents pitch contour in speech' (1967:442). Sakai, also a linguist, says, '*warabeuta* is the most naturally tuned in singing Japanese songs, connected with words, and it can be said that *warabeuta* is the origin of songs' (1998:65), and '… *warabeuta* is created by children and transcribed by adults thereafter' (ibid.:67). Along with the same lines, the musicologist Koizumi advocates the close relationship between speech rhythm and musical rhythm: '… the isochrony of the Japanese syllable [i.e. the near-identity of duration across moras – HN] in careful speech can be observed in *warabeuta*, folk songs and ballads music' (Koizumi 1977).

These children's songs are often performed with physical movements such as clapping and/or marching in time along with the music. The importance of kinetic involvement for acquisition of L1 speech timing has been discussed in various places (e.g. Gardner 1993). Washio (1997) discusses the close interaction between physical movement and children's music based on British nursery rhymes. She argues that stress accentuation in English plays an important role in developing the English bouncing rhythm. That is, in English poetry, the meter is created as a consequence of foot repetition, in which the combination of strong/weak stress (either iambic, trochaic, or even anapaestic or dactylic unit) are represented as one unit. According to Washio, these strong/weak alternations create an effect of limping or bouncing rhythm which can be observed in English nursery rhymes.

Japanese *warabeuta* also seem to emphasise rhythm rather than pitch/melody movement. Here is an example. The following play-song is well known for having no specific pitch required.

(4)

Lyrics are given and children usually sing (or chant) the song accompanied by a sort of complicated demonstration of the contents of a lunch box. The example shows repetition of a basic rhythm. The melodic pattern can be varied. Some studies present evidence of acquiring rhythmic patterns of their language by infants as early as five month old (Nazzi & Ramus 2003, Nazzi & Jusczyk 1999).

This suggests that spontaneous singing by children in a natural setting can display essential, unsophisticated rhythmic qualities, possibly with an unspecified melody line (or intonation in speech). Therefore we argue that it is appropriate to adopt children's songs as data to examine the possible rhythmic interaction between Japanese words and notes, where the high degree of spontaneity might invoke the behaviour of rhythm unit in speech with constraints of rhythm unit in music.

3.1 Materials

30 songs were collected from several song books featuring children's traditional folk music (e.g. *101 Favourite Japanese Songs* 1983). 30 additional songs were transcribed by the author from a commercialised tape entitled *Dooyoo Daizenshuu* ('A great compilation of children's songs' 1991), making a total of 60 songs. All the songs were divided into three categories in 20 songs for each group.

The first group (hereafter, G1) consists of *warabe-uta, asobi-uta*, Japanese folk songs and old airs of Japan, in total 20 songs. *Warabe-uta* ('children's songs': *warabe* is an old term for children, and *uta* generally refers to song even now) and *asobi-uta* ('play songs') are considered to be a genre of traditional folk songs which are spontaneously created for (and sometimes by) children. The songs have survived for over 100 years. According to Washio (1997), the basic style of *warabeuta* was established in the Edo period (1603-1867). The first compilation was published around 1820 (ibid.:20), and generally considered as 1) created in natural settings 2) mora-based and 3) reflecting Japanese speech rhythms. Koizumi (1977) says, '*Warabeuta* is strictly related to the theory of rhythm and melody, and enables us to probe the root of Japanese sense of rhythm' (1977:91). The following is an example phrase of this song category.

(5) "*Kagome kagome*" ('Bow down')
– originated in Chiba prefecture.

The lyrics mean 'Bow down, *Kagome*, a bird in the cage', and this song is sung with a guessing game. As seen here, only four notes are used to compose the first and thematic phrase. This traditional scale is maintained throughout the song. It is also apparent that one note (a crotchet, a quaver or a semiquaver) bears a mora, which supports the "one-mora per note" principle.

The second song group (G2) includes 20 songs from 10 *dooyoo* (also children's songs, but composed by professional musicians) and another 10 from *shooka* songs (another genre of children's songs). These songs are not characterised as the same as *warabeuta* in its strict meaning, but worth including as songs which reflect natural speech in terms of maintaining the traditional rhythm patterns of Japanese songs for children. Songs from both subcategories are adopted under the criterion that they are composed based on the structures of traditional Japanese music such as the pentatonic scale[4] and *miyakobushi*[5]. Despite *dooyoo* and *shooka* songs being composed by professional song writers, I have accorded this group a position of newer but quite similar characteristics to *warabeuta*. An example is given below.

(6) "*Kamome no suihei-san*" ('Seagulls, the sailormen')
– composed in 1933.

The text means 'Seagulls, looking like sailormen'. This song does not require kinetic activities, but a traditional five-note scale is maintained

[4] The pentatonic scale is defined as '… any of several scales consisting of five notes, … the 1st, 2nd, 3rd, 5th, and 6th degrees of the major diatonic scale' (Collins English Dictionary 2000).
[5] *Miyakobushi* is the minor-modal version of the pentatonic scale.

throughout. Although the majority of moras are mapped on to individual notes, the temporal values vary from a semiquaver to a dotted-crotchet; the ratio is 1:6.

Finally, for the third group (G3), another 20 songs are adopted from 10 *dooyoo* and 10 *shooka*. Songs of this group are free from the constraints of the tonality of the pentatonic scale. As discussed above, *dooyoo & shooka* songs are not as old as *warabeuta*, but were created before 1945 when the second Westernisation occurred and have survived over half the century. Some songs are excluded for obvious reasons such as, 1) if they start with *auftakt* ('upbeat') and the composer is unknown, then the music probably reflects aspects of German songs. According to Koizumi (1977), *auftakt* is an archetypal feature of German style and rarely seen in Japanese (traditional) songs; 2) lyrics are written in a literary style, not in a colloquial style. The (only) difference between this group and G2 is that the songs in G3 are not constrained to preserve the pentatonic contour. Therefore, the notes F and B are seen in a C major song (where the pentatonic scale would allow only C, D, E, G, and A), and this makes the songs sound relatively modern and Westernised. Let us look at an example below.

(7) "*Shoojooji no tanuki-bayashi*" ('Racoons' drum-beating at *Shoojooji* temple') – composed in 1924.

The rough meaning of the song text is "Come join us at the moon lit garden of *Shoojooji* temple". Although this song was composed in an earlier period than G2, it is categorized as G3 (more modern music) due to the employment of the Western scale. (The scale here is pentatonic, but an F appears later in the song). Some two-mora mappings are observed in the first, second, and seventh bars (e.g. *shoo* on either a crotchet or a quaver, and *min* on a dotted quaver).

The feature held in common by all 60 songs is their *familiarity*. For example, the selection of *shooka* can be justified in that nobody would need sheet music to sing them accurately. These songs are widely sung though few people (probably only singers) are concerned about the relationship of notes and lyrics. Many naïve performers cannot read music, nor are they aware of the concept of the mora. We do not think about notes, words, rhythm, melody movement, tempo, or other structural factors when we sing songs. This can be one of the reasons to investigate the rhythmic relationship of lyrics (words in song, or timing in speech) and notes (timing in music), where some subconscious interaction of speech components and musical factors may take place.

My null-hypothesis is thus that the relationship between moras and notes is identical in all three song types, i.e. all songs are written (or transcribed) for children, and all which are composed before 1945 when the second Westernisation occurred. In Tanaka's study, the period of composition and the type of songs were correlated with rhythmic changes, which in turn were related to morphological factors (Japanese native vs. loan words, etc.) and mora-structural differences (S-moras vs. NS-moras).

However, we might see some different patterns of interaction between mora and notes according to the delineated categorisation. First, G1 *warabeuta* consists of more 2/4 measures songs than any other group due to having been composed in an older period and of being sung with kinetic games. Secondly, the musical structure (such as the pentatonic scale) might yield a difference between newer song groups: G2 and G3, even if songs in both groups are composed in the same period. In order to compare our results with those of Tanaka, we shall examine to what extent our songs exhibit a one-mora/one-beat structure, with the expectation that newer (and especially non-pentatonic) songs will show a shift towards a many-mora-to-one-note distribution.

3.2 Analysis technique

We have tabulated the number of mora assigned to each note type. All phonological sequences were divided into 12 groups, i.e. one-mora: CV, C, V, N, two-moras CVCV, CVC, CV1V1, CV1V2, CVN, V1V1, V1V2 (all abbreviated (C)Vμ), and three-moras. As for note types, the following eight symbols are used in the calculation: SQ (𝅘𝅥𝅯)– semiquaver (sixteenth note), DSQ (𝅘𝅥𝅯.) – dotted semiquaver, Q (♪) – quaver (eighth note), DQ (♪.) – dotted quaver, C (♩) – crotchet (quarter note), DC (♩.) – dotted crotchet, M (𝅗𝅥) – minim (half note), and DM (𝅗𝅥.) – dotted minim.

4. Results

The overall results are shown as in Table 3 from the total of 60 songs. There were 2852 moras in all, and each column indicates the percentage of each mora type to this figure.

Table 3 The ratio of mora vs. notes assignment

	SQ	DSQ	Q	DQ	C	DC	M	DM	Total
CV	**11**	0	**32**	**8**	**17**	1	2	1	**72%**
C	0	0	0	0	0	0	0	0	**1%**
V	2	0	**6**	2	2	0	0	0	**12%**
N	0	0	1	0	1	0	0	0	**3%**
	SQ	DSQ	Q	DQ	C	DC	M	DM	
CVCV	0	0	**1**	0	0	0	0	0	**1%**
CVC	0	0	**1**	**1**	**2**	0	0	0	**4%**
CV1V1	0	0	0	0	**1**	0	0	0	**2%**
CV1V2	0	0	**1**	0	**1**	0	0	0	**2%**
CVN	0	0	0	0	**2**	0	0	0	**4%**
V1V1	0	0	0	0	0	0	0	0	**0%**
V1V2	0	0	0	0	0	0	0	0	**0%**
3-moras	0	0	0	0	0	0	0	0	**0%**
Total[6]	**14%**	**0%**	**42%**	**12%**	**26%**	**2%**	**2%**	**1%**	**100%**

Figures in bold and highlighted are those higher than 5%. All of them are in the one-mora section. More than one-third of all moras were S-moras (CVs and Vs altogether) mapped onto quavers (32+6=38%). The second highest numbers are also S-moras and mapped onto crotchets (17+2=19%), followed by on semiquavers (11+2=13%). All these are relatively regular timing in terms of having no dots which make the nesting note half times longer and often generate some limping effect in rhythm. On the contrary, two mora mapping is realised at much lower rate. Almost all the cells resulted in holding zero percentage except for a few combinations such as CVC mapping onto crotchet. We will later examine the internal distribution for the two mora mapping which should be considered separately from one mora mapping.

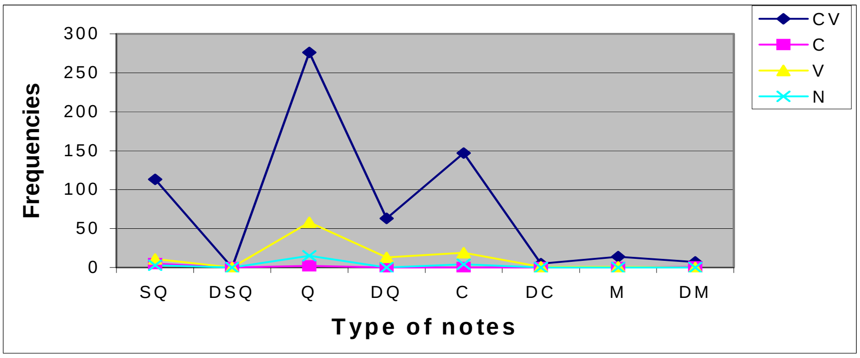
The following sections will show individual results from each song group (G1~G3), first for one mora realisation from 4.1 to 4.4, then for two mora realisation from 4.5 to 4.8. The argument starts from observations of

[6] Each total figure (%) does not match the calculation at glance, as this is all from the results of rounding up/down.

results from each group, and it is developed into comparison of each group, 4.4 for the one mora and 4.8 for the two mora.

4.1 One-mora realisation in G1: warabeuta, etc.

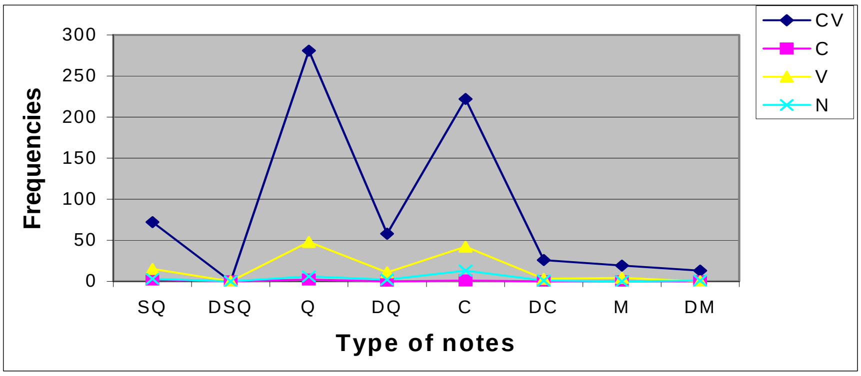
Fig. 1 One-mora assignment on notes in G1(1-20) songs



It is clearly observable that majority of S-moras are mapped onto quavers, followed by crotchets and semiquavers. The second major relationship can be seen for single-vowel moras, which map onto quavers first, then crotchets and semiquavers.

4.2 One-mora realisation in G2: dooyoo & shooka (pentatonic)

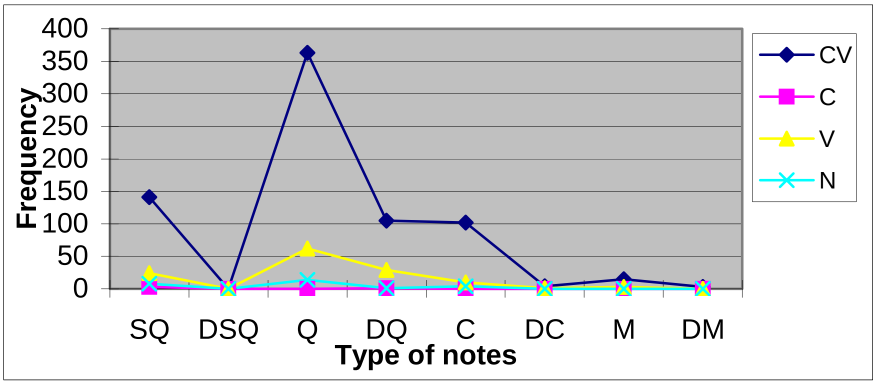
Fig 2. One-mora assignment on notes in G2(21-40) songs



The pattern of S-moras assignment is very similar to that of G1 (Fig. 1), but the ratio of mapping onto crotchets seems to be higher than the results of G1. This may be explained by fact that there are no game-songs in this group but only songs for singing, hence 2/4 measures are rarely seen. This might be from the reason that one-beat (on which 1-mora unit is mapped) is generally represented as a crotchet rather than a quaver. We will see if the similar phenomenon is observable in the G3 songs.

4.3 One-mora realisation in G3: shooka & dooyoo (non-pentatonic)

Fig. 3 One-mora assignment on notes in G3(41-60) songs



The first noticeable thing is the peak for S-moras on quavers, which is even more marked than in the previous two groups – the actual number of occurrences is (G1:351 / G2:337 / G3:438). Second, the one-mora mapping rate of S-moras on crotchets is extremely low compared to that of G1 and G2. The mapping onto crotchets in G3 is even less than that onto semiquavers. We will examine this phenomenon with numeric information of the data in the next section, where the comparison amongst all the song groups is discussed.

4.4 Comparison of one mora representation of all groups

In this section, we compare the internal relationships of only 1-mora/1-note interaction for each group, separately from 2-moras per 1-note results which will be shown in the later section. First the number of one mora mapping for each group is displayed in Table 4.

The main concern of this one-mora comparison is to examine the difference between S-moras and NS-moras for interaction with each note type, so all the NS-moras were integrated as one category appeared as NS-moras in the column for the table. Numbers are converted into percentile in order to make the comparison easier.

Table 4 The comparison of 1-mora/1-note relationship for Song G1~G3

(G1)

	SQ	Q	DQ	C	Total
S-moras	16	38	9	20	**82%**
NS-moras	3	10	2	3	**18%**
Total	**18%**	**48%**	**10%**	**23%**	**100%**

(G2)

	SQ	Q	DQ	C	Total
S-moras	9	36	7	29	**81%**
NS-moras	3	7	2	7	**19%**
Total	**12%**	**43%**	**9%**	**36%**	**100%**

(G3)

	SQ	Q	DQ	C	Total
S-moras	16	42	12	12	**82%**
NS-moras	4	9	4	2	**18%**
Total	**20%**	**51%**	**16%**	**13%**	**100%**

As we can see above, the total ratio of S-moras and NS-moras are nearly the same for all groups (approx. S-moras: 80% vs. NS-moras: 20% in all three groups). However, the allotment to musical notes (as total) does not seem to show any uniformity. For example, the assignment of single moras (the total of S-moras and NS-moras) to SQ and Q in G1 and G3 are quite similar (SQ/Q in G1 = 18/48 = approx. 0.38; SQ/Q in G3 = 20/51 = approx. 0.39), but as observed above, G2 does not seem to share this proportion (SQ/Q in G2 = 12/43 = approx. 0.28). This irregularity can be applied to the assignment on crotchets (C) and dotted crotchets (DC). Assignment of single moras to crotchets differs considerably across groups (G1:23 / G2:36 / G3:13). G1 and G2 share the similar percentage figures for DC assignment (10% and 9%, respectively), whereas G3 does not accord with their results (16%).

In other respects, no particular consistency is observed in the results. Another inconsistent distribution is the assignment to crotchet (C) for the S-mora type (S-mora/C combination). The assignment to crotchet is much less in G3 (12%) than G1 or G2 (20% and 29%, respectively). Moreover, if we compare the figure with the S-moras/Q combination, the ratio of G3

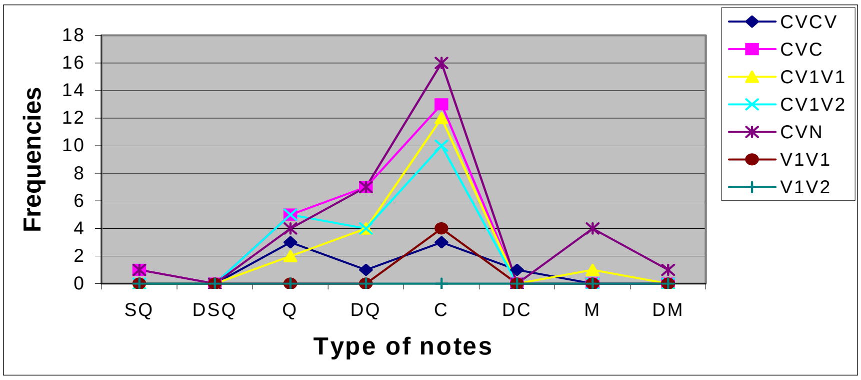
(C/Q of S-mora in G3 = 12/42 = 3.5) is much greater than the ratio of G1 and G2 (1.9 in G1 and 0.8 in G2).

The number of NS-moras is also fluctuated and indicates no particular consistency within/between each groups. The ratio of NS-moras to S-moras also appears to show no particular similarities between each group, except the similar rate of crotchets in G1 and G3 (3/20 = approx. 0.15 in G1 and 2/12 = 0.17 in G3, whereas 7/29 = approx. 0.24 in G2). A X^2 (Pearson Chi-Square = 139.037, DF = 14, P = 0.000; p<.01) shows that the discrepancy is statistically valid over the three matrices, despite the apparently standardised figures in general (e.g. S-moras: 80% vs. NS-moras: 20% in the Total columns).

The next four sections will see the results in two mora assignment. All the data for 2-moras patterns show raw figures because the small numbers of percentiles did not allow for reliable statistical calculations. Almost all figures are less than five and many are zero.

4.5 Two-mora realisation in G1: warabeuta, etc.

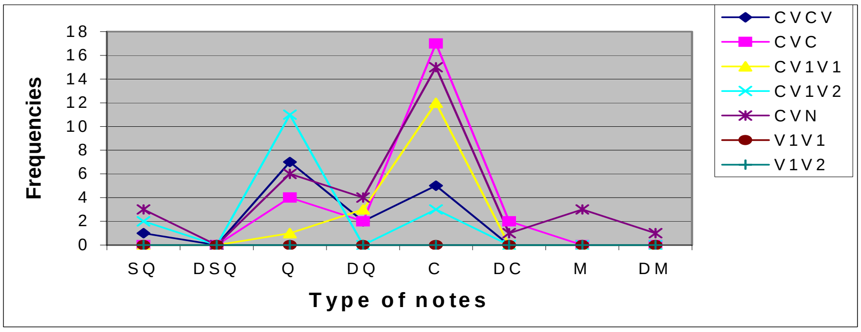
Fig. 4 Two-mora assignment on notes in G1(1-21) songs



Monosyllabic 2-mora sequences are predominantly realised as crotchets. This is noteworthy considering the large number of 1-mora units mapping onto quavers demonstrated in Fig. 1. Since a crotchet is twice as long as a quaver, it can support longer units (CVμ, such as CVC, CVV or CVN) naturally, while a quaver has less durational support for these extended moras which have something extra in addition to the basic CV or V. Among others, CVN assignment to a single note is the most frequent pattern followed by CVC assignment. The lengthened vowels (CV1V1 and CV1V2) also appear prominently.

4.6 Two-mora realisation in G2: shooka & dooyoo (pentatonic)

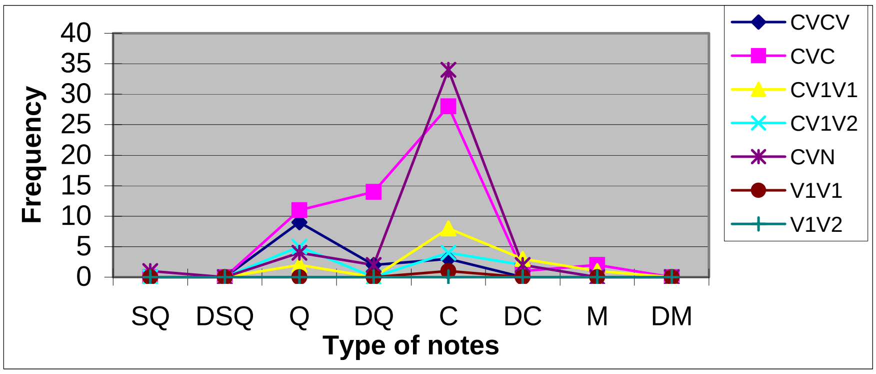
Fig. 5 Two-mora assignment on notes in G2(21-40) songs



The other results are similar to those for G1 songs: two-moras sequences are realised as crotchets, although we do *not* observe a distinctive peak at the crotchet (or only a very weak one), as identical with the results in G1. This supports the generalisation that one-mora units map onto quavers while 2-moras sequences tend to be realised as crotchets. On the other hand, we see some differences in two-mora behaviour between G1 and G2. Some two-mora types (such as CV1V2 and CVCV) are realised as quavers greater than as crotchets in this song group. This phenomenon is not observed in G1, and needs further analysis.

4.7 Two-mora realisation in G3: shooka & dooyoo (non-pentatonic)

Fig. 6 Two-mora assignment on notes in G3(41-60) songs



Here the number of tokens reached as high as 34 which is twice as much as the values in the other two groups. (16 for G1 and 17 for G2). We are led to suspect that there is a difference in the number of moras in total in this type of music. We also see an unusual development here; while the overall tendency is for two-mora sequences to be realised as crotchets, a considerable number of CVC sequences are mapped into dotted quavers, suggesting a possible value of one-and-a-half rather than two moras. However, it seems that only CVNs and CVCs distinctively reach the high peak and the frequency of other combination are quite low. We will subject this impressionistic observation to numeric/statistic examination in the next section.

4.8 Comparison of two mora representation of all groups

This section will discuss the overall similarities and differences of G1~3, based on the actual (raw) figures obtained from the tabulation of 2-moras units assigned to one note.

Table 5 The number of 2-moras units assigned to single notes in G1~G3

G1

	SQ	Q	DQ	C	Total
CVCV	0	3	1	3	7
CVC	1	5	7	13	**26**
CV1V1	0	2	4	12	33
CV1V2	0	5	4	10	19
CVN	1	4	7	16	**28**
Total	2	19	23	54	113
					211

G2

	SQ	Q	DQ	C	Total
CVCV	1	7	2	5	15
CVC	0	4	2	17	**23**
CV1V1	0	1	3	12	16
CV1V2	2	11	0	3	16
CVN	3	6	4	15	**28**
Total	6	29	11	52	98
					196

G3

	SQ	Q	DQ	C	Total
CVCV	0	9	2	3	14
CVC	0	11	14	28	**53**
CV1V1	0	2	0	8	10
CV1V2	0	5	0	4	9
CVN	1	4	2	34	**41**
Total	1	31	18	77	127
					254

The total figures for distribution of CVCs and CVNs are bolded, as these two types achieved the highest mapping in all three groups. These two mora types will be then examined first. The total number of CVCs in G3 takes up more than double as the other groups (G1:26, G2:23, and G3:53); CVNs are mapped likewise (G1:28, G2:28, and G3:41) although the difference is not as high as CVCs. Therefore it seems G3 songs bear many more bimoraic units onto one note than other two song groups, and leads to a suggestion that NS-moras tend to form a bimoraic unit and behave as a part of the unit: μ of CVμ in lyrics of newer songs.

However, if we compare the grand total numbers of each group, the results are not as different as these two mora types (G1:211, G2:196, and G3:254). In addition, the subtotals of quivers and crotchets show interesting results; i.e. the mapping of 2-moras on crotchets are much greater than on quavers in all groups (G1:19/G2:29/G3:31 for quavers, whereas G1:54/G2:52/G3:77 for crotchets). This indicates the crucial factor of rhythmic interaction between speech unit and musical unit. As for CVCVs, G2 and G3 resulted in almost the same numbers in total (G2:15, G3:14), but the internal details differ. In G3, nine CVCVs out of 14 are assigned to quavers, which cannot be ignored if we consider the durational differences between CVCV and CVμ. CVs are generally considered as rather independent units than NS-moras (μs) (see Han 1992, Homma 1981). Therefore the bimoraic combination: CVCVs bear longer duration than CVμs. In this sense, mapping CVCV onto the relatively short note: quaver nine times in G3 is quite interesting compared to other ratios of this assignment in G1 and G2.

To test the significance of the high ratio of particular combinations (such as CVCs/CVNs onto crotchets) and the low assignments (any mora types mapping onto semiquavers which often resulted in zero), SAS (exact methods to calculate the p-value) was conducted, and obtained results were highly significant (Pr >=ChiSq 0.0013; $p<0.001$).[7] The value is significant enough to see the discrepancy amongst mora types (CVC, CVN, etc.), note types (SQ, Q, etc.) vs. song groups (G1~G3). The test results leads to the validity of the propensity of distributing two types of CVμs to certain notes, which further yields one possible interpretation that bimoraic unit can be compressed and assigned to one musical unit in newer songs.

5. Discussion

We have observed a set of data on what type of mora is produced using what value musical unit, using Japanese children's songs. The results looked at two different distributions: 1-mora assignment and 2-mora assignment. The generalisation will therefore start from the 1-mora distribution.

5.1 Generalisation of 1-mora mapping

As observed in the results section (4.4), the interaction of 1-mora mapping with each note type did not hold the consistency among the three groups,

[7] The test method differs from that of 1-mora distribution, due to the very small value of distribution for calculation, especially in the SQ columns of which many figures are zero.

despite the uniform ratio of S-moras and NS-moras in the grand total (approx. 80% and 20%, respectively). The discrepancy is obvious from the figures in Table 4, where the percentage shows no particular sign of similarities between each group. This was statistically valid.

The only consistency is observed when we compare the raw figure of mapping 1-moras in G3 to that of G1 and G2. In G3, an allocation of 1-mora units to crotchets was extremely low. This might be an indication of a current preference for mapping onto shorter notes, or it might even suggest that the duration of the single mora is getting shorter. This prominent difference needs to be examined in accordance with analysis of the data from two-mora interaction with notes.

5.2 Generalisation of 2-mora mapping

In general, 2-mora mapping was quite low in all (13% of all moras), but the internal distribution shows that CVCs and CVNs are the two major 2-mora types which are realised onto one note. This may support the general argument of the low independence of NS-moras, especially CVCs. Kubozono (1999) indeed postulates that the first member of geminates: Q is particularly dependent and tends to behave as a part of bimoraic unit: CVC, rather than an independent unit. However in the present study, it is observed that a moraic nasal: N has the least independence and a unit of CVN is often observed. This can be related to the fact that the nasal is often realised as nasalisation on the vowel rather than an independently articulated unit.

Both CVCs and CVNs are outstanding in terms of their tendency to map onto a single note throughout all song groups. This tendency is prominent particularly in newer style of music. This result accords with Tanaka's findings, in which data is collected from a much more recent period. The results lead to the assumption that there may be a possibility of a closer relationship between the concept of syllable and the mora in Japanese, *if* the unit is bimoraic or monosyllabic. In this assumption, the two-mora sequences (CVμ) can be a monosyllabic unit, which were observed in this study as single speech unit represented onto a musical note.

Another possible reason for the high ratio of 2-mora mappings onto one note in this group may lie in the use of loan words. As Tanaka points out, modern (or newer) songs tend to employ more loan words than traditional songs. The syllable structure of English, as the major source of loan words, might influence the constituents of the lexicon of lyrics (e.g. yielding more NS-moras), and consequently exert an effect in compressing two moras in Japanese into one syllable unit onto one note, thus, behaving like a single speech unit.

6. Conclusion

The results from the measurement of speech units and musical units in 60 children's songs generally support the summary of Tanaka's analysis; i.e. the interplay of both units with each other creates the shift towards the propensity of having more moras per note (or fewer notes per mora) in new songs. Almost the doubled number of bimoraic tokens is obtained in G3 than that of G1 and G2, which were obtained from the observation of two-mora mapping.

The present study also demonstrates the importance of considering the time value of music notes. For example, the predominant allocation of monosyllabic (bimoraic) units were mapped onto crotchets through all the groups as seen in Table 5, in contrast to the singleton mora mapping onto quavers. This suggests that the musical duration of notes affects how many (and what type of) speech units can be held, leading to the further speculation of why a considerable number of CVC and CVN sequences are mapped onto dotted quavers.

The consideration of musical timing also demonstrates the tendency of one-mora unit (CV) mapping onto shorter notes (quaver rather than crotchets) in newer non-traditional songs, which can be another piece of evidence that a speech unit is compressed in a shorter musical phrase in modern music. The two-fold approach (one-mora and two-moras mapping) both suggests that the bimoraic (or monosyllabic) unit is a possible timing unit in Japanese.

A future related study could compare the results of this study with more data from modern (pop) music. It would be enlightening to compare the traditional songs to very modern songs for children in which a great deal of influence (of Western and/or even non-Western music such as Indian, or other folk music which have different style, beat, tempo and structures in general) is possibly in effect in composing and singing. Alternatively, some experiments can be conducted on native speakers to see how they match a speech phrase to a musical phrase. The test will be given to groups from different generations, in order to discover which note values are thought to best suit speech units composed of different numbers of moras.

References

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Collins English Dictionary – 21st century edition. (2000).

Dan, I. (1999). *Watashi no nihon ongakushi* (*History of Japanese music from my perspective*). Tokyo: NHK Library

Dooyoo Daizenshuu (1991). A compilation of children's songs. Apolon. APTF-5088 *edition*. Fontana Press.

Han, M. S. (1992). The timing control of geminate and single stop consonants in Japanese: a challenge for non-native speakers. *Phonetica* **49**. 102-127.

Homma, Y. (1981). Durational relationship between Japanese stops and vowels. *Journal of Phonetics* **9**. 273-281.

Kindaichi, H. (1967). *Nihongo On'in no Kenkyuu* (*The study of Japanese phonology*). Tookyoodoo.

Koizumi, A. (1994). *Ongaku no kongen ni arumono* (*Something fundamental of music*). Heibonsha Library.

Koizumi, A. (1977). *Nihon no oto* (*Sounds of Japan*). Heibonsha Library.

Kubozono, H. (1999). Mora and syllable. In Tsujimura, N. (ed.) *The handbook of Japanese linguistics*. Oxford: Blackwell. 31-61.

Nazzi, T. & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication* **41**. 233-243

Nazzi, T. & Jusczyk, P. W. (1999). Five-month-old infants' discrimination of languages. *Proceedings of the Annual Boston University Conference on Language Development* **23**. 519-528.

Palmer, C. & Kelly, M. H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language* **31**. 525-542.

Patel, A. D. & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition* **87**. B35-B45.

Risset, J. C. (1991). Speech and music combined: an overview. In Sundberg, J., Nord, L. & Carlson R. (eds.) *Music, language, speech and brain: proceedings of an international symposium at a Wenner-Gren Center, Stockholm 1990*. MacMillan Press (Scientific & Medical).

Sakai, Y. (1998). Acoustic characteristics of a special mora 'sokuon' in Japanese songs. *Journal of the Phonetic Society of Japan* **2**. 63-71.

Selkirk, E. (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge: MIT Press.

Tanaka, S. (2000). Nihon kayookyoku ni okeru rizumu no hensen (Rhythmic transition in Japanese pop music). *Proceeding of the 14th annual conference of phonetic society of Japan*. 153-158.

The Japan Times (ed.). (1983). *101 Favourite Japanese Songs Taught in Japanese Schools*.

Ujihira, A. (1996). Kashoo ni miru nihongo no tokusho moora. (Japanese non-standard moras observed in songs). Japanese Society of Phonology (ed.). *Study of phonology – theory and practice*. Kaitakusha. 71-76.

Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany: State University of New York Press.

Washio, N. (1997). *Warabeuta to naasarii raimu – Nihongo to eigo no hikaku rizumu koo* (*Warabeuta and nursery rhyme – Comparative study of speech rhythm in Japanese and English*). Bansuisha.

Zen-on Music Publishers. (ed). (1999). *Purofessyonaru yuusu: kayookyoku no subete besuto 1100* (*Professional use: best songs of all pop music 1100*).