

Sound Symbolism in English and Japanese: A Comparative Analysis of the Expressions Used in Comic Books

Yoshiharu KUMAGAI

1. Introduction

This study explicates where and how sound-symbolic properties may be realized in the expressions of sound-emitting activities in English and Japanese. Special focus is placed on the sound-symbolic relations between the phonological properties of individual segments and larger components inside a syllable, and the magnitude and speed of relevant sound-emitting events. This study aims at clarifying differences and similarities in the manner of linguistic realization of sound symbolism in English and Japanese by analyzing the comparable data of the two languages. It is hypothesized that while sound symbolism in English can mainly be realized in the brightness of the nuclear vowel inside the accented syllable of the words that describe sound emitter and/or sound-emission events, it is more likely to be found in the voiced/voiceless contrast of the consonant of the relevant Japanese words, although English may utilize other options available in Japanese and vice versa.

To test the hypothesis, this study examines the data from the Japanese comic books and their translated version, where the same sound-emission activities in the story are expressed in each of the two languages. Specifically, by analyzing the activities of card-grabbing competitions developed in *Hyakunin Isshu* and described in the Japanese best-selling comic books, the traits of sound symbolism envisaged in the English and Japanese expressions are sampled and compared in terms of visual information and expressions used in each scene.

The structure of this article is as follows. The next section reviews some phonological and sound-symbolic properties of English and Japanese relevant to the present study. It also addresses past case studies conducted by

the present author that point to the correspondence between vowel brightness and sound emitters' physical properties in English. In this study, the findings in the past works are extended to sound-emission events—specifically, the sound-emission activities that accompany the snatching, swiping, and/or bumping cards on the mat in the traditional Japanese game. The third section describes the details of the analytic method employed in this study for exploring the loci and manners of the realization of sound symbolism in English and Japanese data, and based on the results of the analysis, argues for the validity of the hypothesis proposed in this study. The final section summarizes the findings of the present study and addresses some issues for further research.

2. Phonological Properties of English and Japanese Relevant to Sound Symbolism

Sound symbolism is assumed to be codified in both English and Japanese in such a way that it conforms to the phonological structure of each of the two languages. The previous studies conducted by the present author (Kumagai, 2015, 2017, 2018, 2019) clarified that many of the English verbs expressing the sound-emission event or activity are likely to take a CVC monosyllabic structure or to add a frequentative suffix (e.g., *-er*, and *-le*) to it, as illustrated in (1) and (2). On the other hand, Japanese seems more likely to depict the manner of sound emission by means of adverbial phrases, as exemplified in (3) and (4). In Japanese, mimetic adverbial expressions tend to have a simple CVCV structure (e.g., *ker*, *ger*, *kusu*) with two light syllables, which may be repeated twice (e.g., *ker-ker*, *ger-ger*, *kusu-kusu*). Thus, although sound symbolism can be realized differently in the two languages, it follows the basic phonological structure of each language. Such expressions are considered idiomatic in the sense that the two languages employ familiar phonological patterns of the language in question.

- (1) Verbs of laughing in English (Kumagai, 2015)
chortle, chuckle, giggle, guffaw, snicker, snigger, titter, teehee, etc.

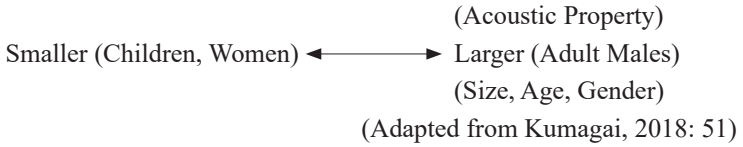
- (2) Verbs of sounds made by birds in English (Kumagai, 2018)
tweet, twitter, caw, quack, gobble, cluck, coo, hoot, etc.
- (3) Laughing in Japanese: Mimetic adverbial expressions and the verb *warau* (“laugh”)
{kera-kera/gera-gera/kusu-kusu/kusutto, etc.} warau
- (4) Sounds made by birds in Japanese: Mimetic adverbial expressions and the verb *naku* (“cry”)
{piyo-piyo/chun-chun/kaa-kaa, etc.} naku

This study tries to identify the locus and manner as well as the degree of realization of the symbolic traits in the two languages based on the copious amount of empirical data. The symbolic properties in the activities or events of sound emission are supposed to be codified in the two languages in different ways. Thus, this study focuses on the acoustic/perceptual properties of the segments utilized in both languages; more specifically, it discusses the properties of “brightness” in vowels and of voicelessness of consonants.

2.1. Vowel Brightness in English

The studies conducted by the present author indicate a strong relation between vowel brightness, which is mainly due to the value of the second formant, and the physical size and/or gender of the sound emitter. It is argued that this is due to English having many vowels (more than Japanese) and to some severe restrictions on the distribution of obstruents inside the syllable in English. Thus, in English, vowels are considered more suitable for expressing sound-symbolic properties. Following the study of Cutler, et al. (1990), this study argues that the following vowel brightness scale (= [5]) is relevant to the expression of sound symbolism in English:

- (5) Properties of Nuclear Vowels in Verbs of Sound Emission
Close, Front (/i:/, /ɪ/) - - - Open, Front (/æ/) - - - Open-Mid to Open,
Back (/ɑ/, /ɔ/, /ʌ/)
Bright ←—————→ Less Bright (Perception)
Higher Frequency of F2 ←————→ Lower Frequency of F2



2.2. Voicelessness of Consonants in Japanese

Unlike English, Japanese does not exhibit complex consonant clusters inside the syllable; in addition, it has less vowels. These facts seem to be related to the possibility that other phonological properties may play a role in sound symbolism inside the syllable of relevant expressions. One such candidate is the voiced/voiceless distinction of the obstruents. Since the restriction on the distribution of obstruents inside the syllable is not as strong as in English, the distinction between voiced and voiceless obstruent may be more easily used to reflect the sound-symbolic characteristics of the sound emitters, as exemplified in (6) and (7), where the voiced expressions (e.g., *boki* and *boko-boko*) are preferred for expressing sound-emitting events involving bigger or denser objects than the voiceless counterparts (e.g., *poki* and *poko-poko*):

- (6) *eda wo {boki/poki}-tto oru* “to break the twig in the {boki/poki} manner”
- (7) *{boko-boko/poko-poko} tataku* “to hit (something) in the {boko-boko/poko-poko} manner”

This difference occurs because, other things (i.e., manner and place of articulation) being equal, the voiceless consonant is considered higher in pitch (Ladefoged, 1982: 168–170) and easily associated with a sound emitter having a small vocal tract.

Thus, it seems reasonable to suppose that in both English and Japanese there is a correspondence between high frequency segments and small sound emitters and between low frequency segments and large sound emitters (Ohala, 1984; Cutler, et al., 1990; Jespersen, 1922, 1933).

2.3. Hypothesis

Based on the analyses developed in Kumagai (2015, 2017, 2018, 2019), the correspondence between phonological properties (i.e., brightness of the nuclear vowel in English and voiced/voiceless opposition of obstruents in Japanese) and sound emitters' physical properties can be extended to the physical properties of sound-emitting activities or events themselves. Thus, in this study, it is hypothesized that while sound symbolism in English can mainly be realized in the brightness of the accented syllable's nuclear vowel in words that describe the manner of the sound emitted, it is more likely to be found in the voiced/voiceless contrast of the obstruents in the relevant Japanese words, although English may have other options that are available in Japanese and vice versa.

3. Data Analysis

3.1. Sources and the Data Examined

This article uses the translated version of the Japanese comic books by Suetsugu (2011, 2012) as analytical data. These works feature a girl who is fascinated by the game *Hyakunin Isshu karuta* and goes through many challenges together with her teammates. In the game, the two competing players try to take the card on which the second half of a traditional Japanese short poem (*tanka*) is written, out of the many laid on the tatami mat as fast as possible. The first half of the poems is read aloud by the speaker, but only the second half is written on the cards in front of the players; thus, the competitors must guess the correct card by matching the contents of the first half, which they hear, and the second half of the poems on the card and snatch the matching one from the mat before the opponent. Currently, *Hyakunin Isshu karuta* tournaments are held at many elementary and secondary schools in Japan, and a national competition is held every year. The competitions generally involve fast and powerful physical movements when the players find the correct card and try to swipe it away from the mat. Since such activities are likely to accompany sound emissions, they are the target of the

analysis in the present article.

Out of 47 volumes published by August 2021, the first two books were translated into English by Varnam-Atkin and Toyozaki, and both the Japanese and English texts can be read in the same copy. All the supposedly relevant tokens involving sound emissions (including but not limited to the activity of snatching cards) were taken from these translated two volumes. For each English token (N = 1,423), its pronunciation, syllable structure, etymological information, and corresponding Japanese original expressions were identified.

Out of the 1,423 tokens, 14 English words (a total of 88 tokens) were found to be relevant to the sound-emitting activities of card-snatching in the game. Interjective expressions (e.g., *yow!*) were excluded from the analysis. The crucial 14 words (including one phrasal example) are listed in (8):

- (8) bang, bash, clank, dong, springing out, swipe, swish, thud, whack, wham, whang, whoosh, whop, zap

The following subsections include the analysis of these 14 words, which will be followed by the analysis of the expressions used in the original Japanese texts where these English words were used as translations.

3.2. Analysis of English Expressions

3.2.1. Overall Tendency

The overall frequency pattern of the relevant words is illustrated in Figure 1. Most of them (13 out of 14) are monosyllabic words, where a disyllabic one (*springing*) involves the *-ing* form. Since the relevant expressions are used to depict instantaneous movements or actions in the game, the use of frequentative suffixes (i.e., *-er*, and *-le*) is unlikely.

In terms of phonological and syllabic structure, as all the examples have a CVC structure, they contain an onset and a coda; this also applies to the stem of the disyllabic word (*spring*). There are several words whose high frequencies stand out: *whack*, *wham*, and *whang* together account for 77.2% of the occurrences of the 14 words. The onset of these words is identical,

and the nuclear vowel is the open front vowel /æ/. It should also be noted that the words beginning with the grapheme *wh* account for as much as 85.2% of the occurrence of the 14 words. Although the onset in question has a variety of pronunciations across English-speaking countries and regions, it is treated here as the voiceless labial-velar fricative (i.e., /ʍ/) because in words beginning with *wh*, the voiceless labial-velar fricative seems to be fairly extensively used across different dialects (Cruttenden, 2014: 234). In contrast, the distribution of the other words is sporadic (i.e., 1 [1.1%] to 2 [2.3%]).

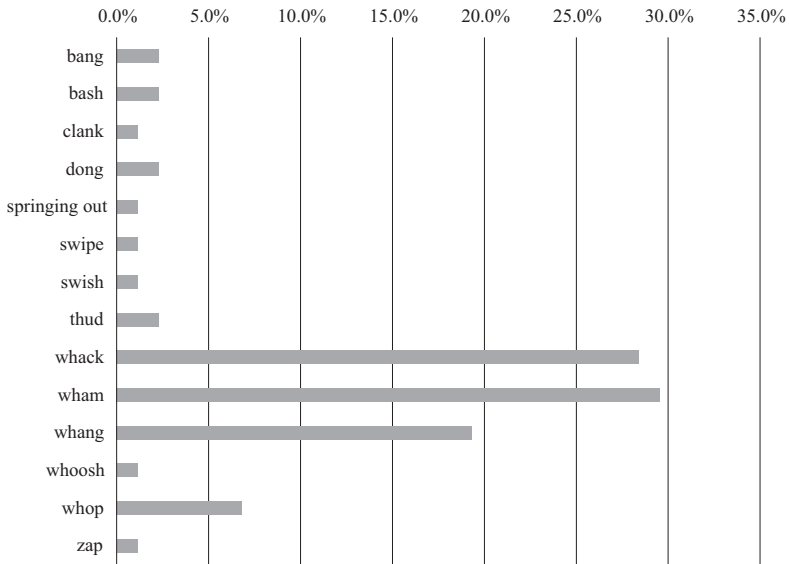


Figure 1. Overall Frequency Pattern of the English Examples

Table 1 illustrates the phonological and etymological properties of the 14 words in more detail. The following characteristics can be observed: first, most of the words contain either open nuclear vowels (10, including the first vocalic element inside the diphthong /aɪ/) or open-mid (1) nuclear vowels, while there are only three close vowels. Ten words have front nuclear vowels,

while four words contain back vowels as a nucleus; etymologically, most of the examples (12 out of 14) are imitative (i.e., derived from the sound relevant to activities and/or movements), while *springing* and *swipe* are non-imitative words.

The etymological information is based on the descriptions in the following electronic dictionaries: *Kenkyusha's English-Japanese dictionary for the general reader with an encyclopedic supplement Version 2*, *Kenkyusha's new English-Japanese dictionary, 6th edn.*, *The Kenkyusha dictionary of English etymology*, *Longman dictionary of contemporary English for advanced learners, 5th edn.*, *Oxford English dictionary, 2nd edn.*, and *Taishukan's unabridged genius English-Japanese dictionary*.

Table 1. Phonological and Etymological Properties of the English Examples

No.	Example	Nucleus	Degree of Aperture	Front/ Back	Imitative Origin?	No. of Syllables	Linear Order of Phonemes	Token	%
1	bang	æ	Open	Front	Yes	1	CVC	2	2.3
2	bash	æ	Open	Front	Yes	1	CVC	2	2.3
3	clank	æ	Open	Front	Yes	1	CVCC	1	1.1
4	dong	ɑ	Open	Back	Yes	1	CVCC	2	2.3
5	spring+ing out	ɪ	Close	Front	No	2+1	CCCVC+VC VVC	1	1.1
6	swipe	aɪ	Open	Front	No	1	CCVVC	1	1.1
7	swish	ɪ	Close	Front	Yes	1	CCVC	1	1.1
8	thud	ʌ	Open-mid	Back	Yes	1	CVC	2	2.3
9	whack	æ	Open	Front	Yes	1	CVC	25	28.4
10	wham	æ	Open	Front	Yes	1	CVC	26	29.5
11	whang	æ	Open	Front	Yes	1	CVC	17	19.3
12	whoosh	u	Close	Back	Yes	1	CVVC	1	1.1
13	whop	ɑ	Open	Back	Yes	1	CVC	6	6.8
14	zap	æ	Open	Front	Yes	1	CVC	1	1.1
								88	100.0

3.2.2. Vowel Brightness in English

Table 1 shows the consistent use of open vowels as a nucleus. There are seven cases of the open front vowel /æ/, with one case of the open back vowel /ɑ/ and the open front vowel /a/. In addition, there is one case of the open-mid back vowel /ʌ/. In 10 out of 14 (71.4%), the nuclear vowels are non-close vowels. Note that the apparent counterexample *spring* is non-imitative etymologically. Although *swipe* is also non-imitative, it contains an open vowel as its first vocalic element. These two examples can be excluded as exceptions, indicating that 10 cases out of 12 involve open and open-mid vowels as the nucleus. These vowels are relevant to a high degree of aperture and less bright vowels with lower F2 than the close front vowel /i/ or the close back vowel /u/. It is remarkable, as will be discussed in detail in 3.4, that non-close vowels are highly likely to be used to describe the forcible action of snatching, grabbing, stomping, swiping, or sweeping away the cards on the tatami mat.

An explanation is in order here for the two cases of close vowels (one for /i/ and one for /u/). The relevant expressions (i.e., *swish* and *whoosh*) are used in the scenes where the card is swiped and immediately flown away in the air. It seems that the English translators pay attention to the sound of the card emitted in the event of it being flown away rather than to the sound of the player forcibly stomping the card when taking it from the tatami mat. In fact, the English expression *swish*, according to the *Oxford English dictionary* (henceforth referred to as *OED*), is an expression describing “a hissing sound like that produced by a switch or similar slender object moved rapidly through the air or an object moving swiftly in contact with water.” Since the card itself is small and thin, the swift sound of it flying away is suitable for an expression with a close front vowel as its nucleus. Likewise, the Japanese expression *pi-shiQ*, with voiceless obstruents (/p/ and /t/ (or /s/)) and close-front vowel in the two syllables, seems to emphasize the swift movement of the card swiped by the player rather than the sound of hitting the card. As for *whoosh*, the nuclear vowel /u/ seems to be used “to emit a dull soft sibilant sound, like that of something rushing through the air; to move rapidly with a rushing sound,” as the *OED* defines it. Thus, the use of close vowels is

considered different from that of open and open-mid vowels that describe the swift and forcible action of snatching (and in many cases swiping) the card. A more detailed discussion on *whoosh*, in comparison with its corresponding Japanese expression, is developed in 3.4.

Further, the open vowel in *whop* and *whang* is used where they depict the sound of simultaneous snatching activities by multiple competitors in games that are happening at the same time. Conversely, in the case of Japanese, a syllable duplication or multiplication is taking place (e.g., *ba-baN*, *ba-baNQ*, and *ba-ba-baNQ*), where the use of the vowel with a lower F2 value than the close front vowel as well as the heavy syllabification in the final syllable is observed, the details of which are included in 3.3.4.

Finally, the word *dong* (whose nucleus /ɑ/ is also an open back vowel) seems to be used for the purpose of depicting a low sound caused by a forcible action because the same word is also used to describe the player falling down on the tatami mat during the game, although this example is excluded from the count as it is not directly related to the action of taking cards. Therefore, it can be said that the open or open-mid vowels in English are consistently utilized by the translators when describing forcible card-swiping activities.

3.2.3. Traits of Sound Symbolism in the Onset

Marchand (1960: 319–334) describes the possible relations between the segment(s) in the onset position and the symbolic properties. Since as much as 85.2% of the words in question contain inside their onset position a voiceless labial-velar fricative /ɱ/ (also represented as /hw/), and since all the words having this segment in the onset are imitative etymologically, it is worth investigating their sound-symbolic property. According to Marchand (1960: 331), the segment /ɱ/ in the syllable-initial position can be used with words “denoting noises of air or breath or forcible movement.” He further notes that the segment is probably used to indicate “the sharp sound of breath at the beginning of a vocal sound or of forcible movement” adding that “it [the segment] is thus somewhat parallel to /h/” (Marchand, 1960: 331).

In its etymological explanation of the verb *whisk*, the *OED* describes that “the spelling with *wh* was adopted as being symbolic (cf. *whip*).” According to *The Kenkyusha dictionary of English etymology* (1997: 1561),

the grapheme *wh* in such words as *whisk* and *whip*, which have undergone various changes in spelling, has a symbolic property, that is, the spelling went through a sort of transformation to reflect the aspirate sound or breath involved in the pronunciation of the onset. According to the *OED* description of the spelling *wh*, it was originally *hw* in such OE-origin words as *hwæt* (= *what*) and, *hwisprian* (= *to whisper*), symbolizing “a voiced bilabial consonant preceded by a breath,” which was developed in two ways. In one direction, the sound in question was changed to a simple voiced consonant /w/; in the other direction, “by the influence of the accompanying breath, the voiced (w) became unvoiced.”

Thus, the segment /ʍ/, which is extensively used as the onset of a syllable of the words *whack*, *wham*, *whang*, and *whop*, seems to be associated with the rapid movement of snatching or sweeping the *karuta* cards by the competitors, arguably with an emphasis on its speed. Other symbolic properties are fulfilled by the nucleus, as far as the relevant examples are concerned.

3.3. Analysis of Japanese Expressions

3.3.1. Overall Tendency

Based on the descriptions of the English examples in Table 1, the corresponding Japanese expressions are listed in Table 2. The transcription in Table 2 utilizes the Romanized orthography of Japanese, with a phonemic transcription added when necessary. For example, *Q* is used for signifying the syllable-final moraic obstruent or Japanese *sokuon*, and *N* for the moraic nasal or Japanese *hatsuon*, for the purpose of reflecting the Japanese original transcription in the comic books as faithfully as possible. Note that the transcription of the syllable-final sequence of *NQ* reflects the orthography of the expression in the original comic works. The examples in Table 2 indicate that the translation into English is not done mechanically on a one-to-one correspondence basis. For example, the English expression *whang* covers a wide range of Japanese expressions (*bashi*, *bashiQ*, *babaN*, *pashi*, *pashiQ*) that contain diverse segments (voiced/voiceless obstruents, close and open vowels, as well as the moraic obstruents and nasals). It is precisely the

purpose of the present analysis to try to explain why such variations may occur rather than having a one-to-one correspondence relation between the Japanese expressions and their English translation.

One of the most conspicuous characteristics of the Japanese expressions in Table 2 is that the expressions are strongly likely to constitute a CVC(C) (heavy syllable) structure in which either the moraic obstruent or moraic nasal closes the monosyllabic words as well as the final syllable of the polysyllabic words. This accounts for 86 cases (97.7%) where only two words (*bashi* and *pasha*, corresponding to English *whang*) are closed by the light syllables. The proportion of monosyllabic and polysyllabic expressions is 60 tokens of monosyllabic expressions (68.2%) and 28 tokens of polysyllabic ones (31.8%). Most of the polysyllabic tokens are disyllabic (27 out of 28 [96.4%]).

Table 2. Phonological Properties of the Japanese Expressions

No.	English Expression	Corresponding Japanese Expression (Token)	%	Initial Consonant of Each Syllable	Linear Order of Phonemes	Vowel
1	bang	doQ (1)	1.1	voiced	CVC	o
		toQ (1)	1.1	voiceless	CVC	o
2	bash	taNQ (2)	2.3	voiceless	CVCC	a
3	clank	gashaN (1)	1.1	voiced	CV+CVC	a + a
4	dong	daN (2)	2.3	voiced	CVC	a
5	springing out	doQ (1)	1.1	voiced	CVC	o
6	swipe	toQ (1)	1.1	voiceless	CVC	o
7	swish	pishiQ (1)	1.1	voiceless + voiceless	CV+CVC	i + i
8	thud	dooQ (1)	1.1	voiced	CV+VC	o + o
		taQ (1)	1.1	voiceless	CVC	a
9	whack	baQ (22)	25.0	voiced	CVC	a
		baNQ (1)	1.1	voiced	CVCC	a
		paQ (2)	2.3	voiceless	CVC	a

Sound Symbolism in English and Japanese

10	wham	baN (9)	10.2	voiced	CVC	a
		baNQ (14)	15.9	voiced	CVCC	a
		paN (1)	1.1	voiceless	CVC	a
		paNQ (2)	2.3	voiceless	CVCC	a
11	whang	bashi (1)	1.1	voiced + voiceless	CV+CV	a + i
		bashiQ (5)	5.7	voiced + voiceless	CV+CVC	a + i
		babaN (2)	2.3	voiced + voiced	CV+CVV	a + a
		pashi (1)	1.1	voiceless + voiceless	CV+CV	a + i
		pashiQ (8)	9.1	voiceless + voiceless	CV+CVC	a + i
12	whoosh	dooQ (1)	1.1	voiced	CV+VC	o + o
13	whop	bababaNQ (1)	1.1	voiced + voiced + voiced	CV+CV +CVCC	a + a + a
		babaNQ (5)	5.7	voiced + voiced	CV+CVCC	a + a
14	zap	pashuQ (1)	1.1	voiceless + voiceless	CV+CVC	a + u
		26 (88)	100.0			

3.3.2. Voiced-Voiceless Opposition of Obstruents in the Japanese Expressions

The obstruents that appeared in the onset position of the monosyllabic tokens (= 60) are /b/ (46), /p/ (5), /d/ (4) and /t/ (5). Voiced plosives (= 50) are more frequently used than voiceless ones (= 10), and bilabial stops (= 51) dominate the distribution. The consonants used in the onset position of the polysyllabic tokens tend to follow either of the following patterns: voiced-voiced (= 9 (*babaN* (2), *bababaNQ* (2), *babaNQ* (5))), voiceless-voiceless (= 11 (*pishiQ* (1), *pashi* (1), *pashiQ* (8), *pashuQ* (1))). There are seven examples of voiced-voiceless patterns (*gashaN* (1), *bashi* (1), *bashiQ* (5)), where the second segment is fricative; in two cases, the second syllable lacks the onset (*dooQ*). In contrast, in no case is the voiceless consonant followed by the voiced one.

Based on the match between visual information and the expression in question, voiceless segments are likely to be utilized for the swift actions of swiping the cards, whereas voiced ones tend to be used for the more forcible and noisy actions. This tendency reflects the validity of the hypothesis, although there are some problematic examples. The details of the result of analysis are presented in 3.4.

3.3.3. Vowels in the Japanese Expressions

The vowels used in the 60 monosyllabic tokens are either open central (= 56 (*taNQ* (2), *daN* (2), *taQ* (1), *baQ* (22), *baNQ* (15), *paQ* (2), *baN* (9), *paN* (1), and *paNQ* (2))) or close-mid back (= 4 (*doQ* (2) and *toQ* (2))). In the case of polysyllabic expressions (= 28), the vowels used inside each of the syllables are either identical throughout: open central (= 9 (*gashaN* (1), *babaN* (2), *babaNQ* (5), and *bababaNQ* (1))), close-mid back (*dooQ* (2)), and close front (*pishiQ* (1)). In some cases, however, the vowels used are different, where the vowel with the higher F2 value is followed by the lower one (= 16 (*bashi* (1), *bashiQ* (5), *pashi* (1), *pashiQ* (8), and *pashuQ* (1))) but not vice versa.

Non-close vowels tend to be used for expressing forcible snatching and/or sweeping and/or bumping cards. Conversely, close front vowels, and possibly close back vowels—as in the second syllable of *pashuQ*—are likely to be utilized for other purposes, such as the speedy actions of snatching cards rather than the noisy bumping of the card or mat. In addition to the voiced/voiceless opposition of obstruents, vowel brightness may also play a role in Japanese. More details are discussed in 3.4.

3.3.4. Duplicative or Multiplicative Light Syllables in Japanese

As mentioned in 3.2.2, light syllables present a few cases of duplication or even multiplication, with the final one closed and made heavy by the dependent mora (*N* or *Q*, or both), as in *babaN* (2 for *whang*), *bababaNQ* (1 for *whop*), and *babaNQ* (5 for *whop*). These expressions are used to describe the simultaneous card-snatching and/or sweeping actions by multiple players participating in the games happening nearby at the same time. The use of heavy syllable expressions seems to be consistent with the English open back nuclear vowel /ɑ/ used in *whop*.

3.4. A Combined Analysis of English and Japanese

The result of matching the relevant expressions (14 English and 26 Japanese expressions, with 88 tokens in total for both languages) with visual information is illustrated in Table 3.

Sound Symbolism in English and Japanese

Table 3. Result of Matching between Relevant Expressions and Visual Information

Example (Token)	<i>wham</i> (26)	<i>whack</i> (25)	<i>wbang</i> (17)	<i>whop</i> (6)	<i>whoosh</i> (1)	<i>bang</i> (2)	<i>bash</i> (2)	<i>clank</i> (1)	<i>dong</i> (2)	<i>springing out</i> (1)	<i>swipe</i> (1)	<i>swish</i> (1)	<i>thud</i> (2)	<i>zap</i> (1)	Total	%
No contradiction between sound symbolic properties and visual information	21 (80.8%)	23 (92.0)	16 (94.1)	5 (94.1)	0 (0.0)	2 (100.0)	1 (50.0)	1 (100.0)	2 (100.0)	1 (100.0)	1 (100.0)	0 (0.0)	2 (100.0)	1 (100.0)	76	86.4
Contradiction	0 (0.0%)	1 (4.0)	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	1 (50.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	0 (0.0)	4	4.5
Lack of enough visual information	5 (19.2%)	1 (4.0)	1 (5.9)	1 (5.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	8	9.1
Corresponding Japanese expression to the above example	<i>baN</i> (9) <i>baNQ</i> (14) <i>paN</i> (1) <i>paNQ</i> (2)	<i>baQ</i> (22) <i>baNQ</i> (1) <i>paQ</i> (2)	<i>bashi</i> (1) <i>bashiQ</i> (5) <i>baban</i> (2) <i>pashi</i> (1) <i>pashiQ</i> (8)	<i>bababNQ</i> (1) <i>bababNQ</i> (5)	<i>dooQ</i> (1) <i>toQ</i> (1)	<i>taNQ</i> (2) <i>gashan</i> (1)	<i>daN</i> (2)	<i>doQ</i> (1)	<i>daQ</i> (1)	<i>toQ</i> (1)	<i>toQ</i> (1)	<i>pishiQ</i> (1)	<i>dooQ</i> (1) <i>taQ</i> (1)	<i>pashuQ</i> (1)	Total	%
No contradiction between sound symbolic properties and visual information	18 (69.2%)	23 (92.0)	8 (47.1)	5 (94.1)	0 (0.0)	1 (50.0)	1 (50.0)	1 (100.0)	2 (100.0)	1 (100.0)	0 (0.0)	0 (0.0)	1 (50.0)	0 (0.0)	61	69.3
Contradiction	3 (11.5%)	1 (4.0)	8 (47.1)	0 (0.0)	1 (100.0)	1 (50.0)	1 (50.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	1 (100.0)	1 (50.0)	1 (100.0)	19	21.6
Lack of enough visual information	5 (19.2%)	1 (4.0)	1 (5.9)	1 (5.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	8	9.1

(14 English and 26 Japanese Expressions, with 88 Tokens in Total for Both Languages)

The data begin with the English words with the *wh*-spelling, which are followed by the others in spelling order, with the corresponding Japanese expressions provided below. In English, 76 cases out of 88 (86.4%) seem to be consistent with the hypothesis, while four cases (4.5%) seem to show contradiction or discrepancy to the hypothesis, and for eight cases (9.1%), the correspondence is not clearly determined because of the lack of enough visual information on the closer investigation into the relation between expression and scene; it turned out that the actions of the players were not clearly or explicitly depicted by the manga and/or with only Japanese captions available.

Compared to the English data, the matching result in Japanese indicates a lower level of consistency. Specifically, while 61 cases out of 88 (69.3%) seem to be consistent with the hypothesis; 19 cases (21.6%) indicate contradiction or discrepancy with the hypothesis; and for eight cases (9.1%), correspondence is not clearly determined because of the lack of enough visual information. If these eight cases are excluded, however, the level of consistency is 95.0% for English and 76.2% for Japanese.

The most striking contradiction or discrepancy with the voicelessness hypothesis in Japanese can be found in the expressions corresponding to the English word *whang*, which means, according to the *OED*, “a resounding blow or stroke, or the sound of such a blow.” Although the images in the comic work depict the swiping activities accompanying sounds, some Japanese expressions utilize the voiceless labial stop (e.g., *pashi* and *pashiQ*) as well as the voiced one in similar scenes (e.g., *bashi* and *bashiQ*). The frequency of these minimal-pair tokens is the same, which means that the voiced/voiceless opposition does not play a role of distinction of meaning.

The existence of duplicative Japanese expressions (*babaN* (2 for *whang*)) mentioned in 3.3.4 suggests that the expressions are amenable to the powerful card-snatching or swiping activities. However, despite the strong tendency to utilize heavy syllables in Japanese expressions, the only two examples of light syllables in the data (i.e., *bashi*, and *pashi*) are also observed in the Japanese expressions corresponding to *whang*; it is not clear why these indicate diverse and conflicting phonological properties (i.e., existence of

minimal-pair tokens, which lack the meaning opposition, existence of a heavy syllable token and of light syllable tokens).

Some other problematic examples deserve mentions here, although the number of tokens is small, and hasty generalization should be avoided. Let us discuss *whoosh* and the corresponding Japanese expression *dooQ*. Both expressions are problematic for our hypothesis, even though this example is found just once. In the case of English, the nucleus is a close back vowel; since it is the least bright one, it is not suitable for the description of strongly swiping the card on the tatami mat. Rather, the English translation seems to pay attention to the sound of swiped cards flying in the air as a result of the powerful action of the player because the word in question is likely “to emit a dull soft sibilant sound, like that of something rushing through the air,” as the *OED* defines it. Conversely, the Japanese expression *dooQ* seems to put more focus on the powerful motion of hitting the card than to the sound of swiped cards scattering together in the air, as indicated by the voiced obstruent /d/, which is contrastive to the consonantal segments in *whoosh*, although use of the close-mid back vowel remains unexplained.

The case of *swish* (one token) is interesting in that both the English expression and the corresponding Japanese one seem to pay attention to the sound of the card being swept away rather than that of bumping it on the mat. The scene occupies nearly half of the page, with the player swiping the card and the Japanese expression *pishiQ* being inserted in the scene using large letters. The use of *swish*, whereby the brightest vowel /i/ is used as a nucleus, seems to focus on the sound of the card flying in the air swiftly; as the *OED* puts it, *swish* describes “a hissing sound like that produced by a switch or similar slender object moved rapidly through the air.” In the case of Japanese, a similar line of argument seems to hold because of the expression *pishiQ*, where the voiceless consonants are used in the onset of the two syllables together with the bright vowels in each of the syllables.

Similarly, the English expression *zap* (1 token) and its corresponding Japanese expression *pa-shuQ* seem to depict the speed and power of swiping the card. It may be the case that the power is reflected in the nucleus of *zap*; the *OED* definition of this word must then be addressed, according to which

zap is “used to represent the sound of a ray gun, laser, bullet.” On the other hand, the Japanese expression *pashuQ* utilizes voiceless consonants in each of the onsets and a close back vowel in the accented heavy syllable. This is not explained straightforwardly, but the use of voiceless segments and of the close vowel in the accented syllable may be related to the swiftness of the card-snatching activity rather than the power exerted when the card is snatched.

Next, a few words are in order for the expression *swipe*, which is not etymologically imitative. Although the scene describes the powerful but stable, confident action of the player snatching the card, the corresponding Japanese expression *toQ* utilizes the voiceless consonant as the onset, although the expression in question is expressed with large-sized letters. It is not clear how this discrepancy can be characterized in Japanese, even though only one example is observed.

Finally, the cases of *thud* (2 tokens) ought to be discussed. The word *thud* is used in two independent scenes: in one scene, a player rushes toward the target card and snatches it, while in the other, another player hits the target card on the mat. In the first case, the use of the Japanese expression *dooQ* is considered consistent with the English expression because of the voiced obstruent in the onset, although the use of the close-mid back vowel remains unexplained, as in the case of *whoosh*. However, in the second case, the expression *taQ* is used, which contradicts the action of bumping the card in the actual scene.

4. Summary and Concluding Remarks

The findings in the present study can be summarized as follows. According to the result of match between the phonological properties of the relevant expressions and visual information, the English data indicate a higher level of consistency than the Japanese ones. In English, the open (/æ/ and /a/) or open-mid (/ʌ/) vowels are consistently utilized when describing the powerful card-swiping activities; this indicates that the aperture of the vowels may indeed be as relevant as the degree of brightness. Additionally, the front-versus-back

tongue position seems to be relevant to the degree of the powerful activity, where the open back vowel /ɑ/ (e.g., *whop*) indicates that a greater degree of energy is involved in the activity compared to the open front vowel /æ/ (e.g., *whack*, *wham*, and *whang*).

The overwhelming distribution of the words beginning with the segment /ʌ/ indicates that some consonants may play a role in signifying sound-symbolic characteristics, as indicated by Marchand (1960). Further, the three expressions *whack*, *wham*, and *whang* are most frequently used in the card-swiping activity; as they share the onset /ʌ/ and the open front nuclear vowel /æ/, these words are minimally distinct by the segment in the coda.

On the other hand, the data in the Japanese expressions seem to indicate that both voiced/voiceless opposition and vowel brightness are playing a role in expressing card-snatching activities, although the role of the former is more conspicuous. Specifically, non-close vowels are likely to be used for expressing the forcible action of snatching and/or sweeping the cards. The close front vowel and possibly the close back vowel are likely to be utilized for expressing the speedy actions of snatching cards or the sound of cards scattering in the air. In addition, the following findings are worth mentioning. Although Japanese generally allows word formation out of a single light syllable, the examples used in the card-snatching activities indicate that the heavy syllable is preferred, where either the moraic obstruent or nasal—or even both—will be attached to the coda of the (final) syllable. The heavy syllable formation by the moraic obstruent or nasal (cf. Kuroda, 1965 (= 1979): 205; Kubozono, 2015: 10) and duplication or multiplication of syllables seem to be relevant to the symbolic properties of expressing powerful and noisy actions; particularly, the multiplication of the same syllables seems to indicate the simultaneous actions by multiple agents, while the use of the voiceless consonant as well as of light syllables may be more amenable to the symbolic properties of the quick, swift actions of snatching and sweeping the cards rather than hitting or bumping them.

This study has also revealed new problems that need to be clarified in the future. There are several cases in which a succession of card-swiping activities is depicted in the series of neighboring comic frames. In such cases,

repeating the identical expressions should generally be avoided. However, if the activities involve similar physical actions in terms of the power and/or speed, how will they be expressed by sound-symbolic words? Can repeating identical expressions be avoided while conforming to their sound-symbolic properties, or will such phonological properties be overridden or superseded, with the result that each of the expressions will be selected arbitrarily? This is an issue to be considered and developed elsewhere. Finally, although the present study focused on the activities of card-snatching, other sound-emission activities found in the same data set ought to be investigated more comprehensively to strengthen the arguments developed in this article.

Acknowledgments

The primary data processing of the two comic books was conducted with the assistance of Miho Tanaka, whom I gratefully acknowledge. This research was supported in part by the Aichi Prefectural University President's Special Research Grant in 2018.

References

- Cruttenden, Alan. 2014. *Gimson's pronunciation of English*, 8th edn. London & New York: Routledge.
- Cutler, Anne, James McQueen & Ken Robinson. 1990. Elizabeth and John: Sound patterns of men's and women's names. *Journal of Linguistics* 26. 471–482.
- Jespersen, Otto. 1922. Sound symbolism. In O. Jespersen, *Language: Its nature, development, and origin*, 396–411. London: Allen and Unwin.
- Jespersen, Otto. 1933. Symbolic value of the vowel i. In O. Jespersen, *Linguistica: Selected papers in English, French and German*, 283–303. Copenhagen: Levin and Munksgaard.
- Kubozono, Haruo. 2015. Introduction to Japanese phonetics and phonology. In H. Kubozono (ed.), *Handbook of Japanese phonetics and phonology*, 1–40. Berlin: De Gruyter Mouton.
- Kumagai, Yoshiharu. 2015. Traits of sound symbolism in English verbs: Evidence from the verbs of laughing. *Mulberry* 65. 25–38.
- Kumagai, Yoshiharu. 2017. Some problems in exploring the traits of sound symbolism: A case study of verbs of crying. *The Journal of the Faculty of Foreign Studies (Language and Literature)*, Aichi Prefectural University 49. 21–34.

- Kumagai, Yoshiharu. 2018. Linguistic and symbolic properties of sound emission verbs: A case study of birds and its theoretical implications. *The Journal of the Faculty of Foreign Studies (Language and Literature)*, Aichi Prefectural University 50. 49–68.
- Kumagai, Yoshiharu. 2019. Light emission verbs in English: An investigation of sound-symbolic and semantic properties. *The Journal of the Faculty of Foreign Studies (Language and Literature)*, Aichi Prefectural University 51. 51–70.
- Kuroda, S.-Y. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: MIT dissertation. [Published in 1979 (Outstanding dissertations in Linguistics), New York & London: Garland Publishing].
- Ladefoged, Peter. 1982. *A course in phonetics*, 2nd edn. New York: Harcourt Brace Jovanovich.
- Marchand, Hans. 1960. *The categories and types of present-day English word formation: A synchronic-diachronic approach*. Wiesbaden: Otto Harrassowitz.
- Ohala, John. J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41, 1–16.

Data Sources

- Suetsugu, Yuki. 2011. *Chihayafuru* (Bilingual Version) Volume 1. Translated by Varnam-Atkin, Stuart, & Yoko Toyozaki (ISBN: 978-4-06-250150-7). Tokyo: Kodansha.
- Suetsugu, Yuki. 2012. *Chihayafuru* (Bilingual Version) Volume 2. Translated by Varnam-Atkin, Stuart, & Yoko Toyozaki (ISBN: 978-4-06-250151-4). Tokyo: Kodansha.

Dictionaries

- Kenkyusha's English-Japanese dictionary for the general reader with an encyclopedic supplement*, Version 2. 2000 (CD-ROM Version, 2000). Tokyo: Kenkyusha.
- Kenkyusha's new English-Japanese dictionary*, 6th edn. 2002 (CD-ROM Version, 2006). Tokyo: Kenkyusha.
- The Kenkyusha dictionary of English etymology*. 1997. Tokyo: Kenkyusha.
- Longman dictionary of contemporary English for advanced learners*, 5th edn. 2009 (DVD-ROM Version 1.1., 2009). Harlow, Essex: Pearson Education Limited.
- Oxford English dictionary*, 2nd edn. 1989. (CD-ROM Version 4.0.0.3, 2009). Oxford: Oxford University Press.
- Taishukan's unabridged genius English-Japanese dictionary*. 2001 (CD-ROM Version, 2002). Tokyo: Taishukan.

直接比較可能なデータの分析に基づく日英語の 音象徴研究

熊谷 吉治

音放出表現が頻出する競技かるたをテーマにした漫画「ちはやふる」の英語訳版から、英語の音放出表現を1,423例抽出した。個々の例に音節構造や核母音、語源の情報、対応するオリジナル版の日本語表現とその音韻的特徴などを関連付けることで、日英語の直接比較が可能な電子資料を作成した。この資料を元に、競技者がかるたを取る際の英語音放出表現(14語(88例))と、対応する日本語表現(26語(88例))に焦点を当て、音韻構造と象徴性との関係を詳しく分析した。

分析から次の事が分かった。日英語の音放出表現は機械的に一対一の対応を示すわけではない。例えば、英語の *whang* は「バシ」、「バシッ」、「ババン」、「バシ」、「バシッ」といった日本語表現が使われる場面に現れている。また、英語では *wh* で始まる語 (*whack*, *wham*, *whang*, *whoosh*, *whop*) が音放出表現として多用されており、これらの語の出現頻度が全体の85.2%を占めている。また、*whack*, *wham*, *whang* の三語で全体の77.2%を占めているが、これらは尾子音のみ異なる単語である。

音象徴性の有力な具現方法や場所として、核母音の明るさや舌の前後位置(=英語)と、子音の有声性(=日本語)が重要な要因になると仮説を立てて検証したところ、仮説との合致率は英語(95.0%)の方が日本語(76.2%)より高かった。加えて、英語では頭子音・尾子音の性質が、日本語では音放出表現内の核母音の明るさと当該表現の重音節化が、音象徴性に一定の役割を果たしうる可能性が浮かび上がった。

さらに、新たな課題も浮かび上がった。カルタを取る力強い(あるいは素早い)動作が漫画のコマの中で連続して現れた場合(例: WHAM/バンッ → THUD/タッ → WHANG/バシ → WHAM/バンッ)、日英語でどのような表現形態を取るのか。同一表現の繰り返しを避ける傾向があるとすれば、音象徴性はどのように担保されるのかを検討する必要がある。また、今回の研究を足がかりに、作成した電子資料に収録された他の音放出表現についても包括的に研究を行うことで、音放出表現内の音素や音節の特徴と音象徴性との関連をさらに検証する必要がある。