

The Usage Mechanism of Japanese Ideophones in the Description of Taste: Morphological and co-occurrence analysis of the description of wines and sakes

Hiroki Fukushima, Mutsumi Imai, Shigenori Tanaka
Keio University, Fujisawa City, Japan
fh@keio.jp
imai@sfc.keio.ac.jp
stanaka@sfc.keio.ac.jp



ABSTRACT: *In this study, we examined ideophones (onomatopoeias or sound symbolism) in taste descriptions of wine and sake using quantitative text analysis and morphological analysis to investigate how sound symbolic expressions function (word count of wine text: 201,294; word count of sake text: 50,147). We found the following co-occurrence tendencies for Japanese ideophones in the tasting corpora: ideophones are used to modify words for taste, rather than words for brewing or grades of wines and sakes. More specifically, ideophones are used to express the sense of ‘appearing’ or ‘finishing’ of flavours (specifically in the Sake Corpus). Morphological analysis of these texts in comparison to the BCCWJ (The Balanced Corpus of Contemporary Written Japanese) revealed that double consonant patterns are a characteristic form in the tasting descriptions. Integrating these findings with those of the text analysis, we concluded that the ideophones are employed to describe breaking points, turning points, and changing processes of taste, rather than stable states. Therefore, ideophones can be distinguished from general adverbs, which mainly describe the states and manners of objects or events.*

Keywords: Ideophones, Onomatopoeias, Sense of taste, Morphology, Co-occurrence analysis, Wine, Sake

Received: 19 February 2017, Revised 10 April 2017, Accepted 3 May 2017

© 2017 DLINE. All Rights Reserved

1. Introduction

Clear and accurate descriptions of taste are difficult to express in Japanese. The main difficulty is likely related to the lack of words for describing taste in the language. Only 10 or fewer Japanese words are used to describe taste directly; words exist for the five basic tastes, astringency, and dryness, as well as for a few complex tastes such as *koku* (‘rich taste’ is the nearest concept). Specific communities such as sommeliers employ more detailed and sophisticated systems of so-called ‘tasting words’. These words can more clearly point out the elements of the taste of wine. However, these tasting words have less power to describe the movements, temporal changes in the mouth, or relationships between taste elements. This difficulty occurs because most tasting words are nouns, while movements, changes, and relationships are probably best described using verbs or adverbs. When commenting on taste, people do tend to use nouns, but verbs and adverbs also play significant roles.

Among adverbs, ideophones can be characterized as vivid sensory words. However, the meanings of ideophones depend upon the domains or contexts of their use. Therefore, we looked at the domain of tasting descriptions for wines and sakes to investigate the roles of Japanese ideophones.

In this study, we attempted to reveal the morphological features and semantic roles of Japanese ideophones in tasting descriptions. Previous studies about Japanese ideophones have mainly focused on their phonology and morphology, with the phonology of ideophones having been studied in domains such as linguistics, cognitive psychology, and cognitive science. The main purpose of examining the phonology of ideophones is to reveal the motivated relationships between sounds and their mental representations. Phonological studies have given some explanation to intuitional sound symbolic phenomena; for example, when comparing *sara-sara* (smooth tactile) to *zara-zara* (rough tactile), the dull sound of /z/ is motivated by tactile friction. The relationships between form and meaning have long been studied in morphological studies of ideophones [7][14]. These forms mainly consist of the numbers of syllables and the patterns of consonants, reduplications, and suffixes.

These previous studies of Japanese ideophones have been based on a strong hypothesis: an ideophone itself has some independent meaning. In this study, we take the perspective that the meaning of an ideophone is determined by its co-occurrence with other words.

1.1 Ideophones

Classifications for sound symbolic words include sound symbolisms, onomatopoeias, and ideophones. In this paper, we use the general term ‘ideophones’ as a common cross-linguistic term for a sound symbolic phenomenon [5][16]. Ideophones can be defined as *marked words that depict sensory imagery* [4]. Japanese ideophones can generally be classified into three categories: phonomimetic words (*gion-go*), phenomimetic words (*gitai-go*), and psychomimetic words (*gijoh-go*). The largest Japanese ideophone dictionary contains more than 4,500 ideophones [12]. However, many of those words are not used consistently in daily Japanese. In this study, we selected 972 ideophones as targets of our investigation. These 972 words are considered to be ‘basic words’ in a concise ideophone dictionary [3].

1.2 Tasting Corpora for Analysis

We used two corpora in this study: a corpus of wine tasting expressions and a corpus of Japanese sake tasting expressions. Each corpus is based on books and magazines written entirely in Japanese, and Table 1 summarizes the details of each. The Data Jacket [13] (the concise abstract of the data) and the references of the Wine and Sake Corpora are shown in the footer¹².

The different measures employed in the study are represented in the first column of Table 1, including types, tokens, sentences, and frequency. Note that ‘Paragraphs’ refers to the number of different wine (or sake) brand descriptions. The second column provides data from the Wine Corpus, based on a wine-reviewing magazine published in Japan (*Real Wine Guide*), while the third

	Wine Corpus	Sake Corpus
Tokens	201,294	50,147
Types	9,449	3,047
Sentences	11,421	2,552
Paragraphs (Brands of Wine)	1,957	1,167
Average Frequency (SD)	10.50 (64.55)	8.47 (34.92)

Note. The paragraphs of each Corpus correspond to different wine brands’ descriptions in the books and magazines. Standard Deviation in parenthesis.

Table 1. Details of the Wine Corpus and Sake Corpus

¹Data Jacket of the Wine and Sake Corpora

Outline of the dataset: Corpora of the tasting descriptions from wine and sake books and magazines written entirely in Japanese. The corpora include both text data of tasting descriptions and detailed data of wines and sakes.

Data sharing policy: Under particular conditions.

column describes data from the Sake Corpus, based on sake-reviewing books and magazines sold in Japan. The sizes of the Wine Corpus and Sake Corpus were 201, 294 words and 50, 147 words, respectively.

1.3 The Importance of Ideophones in Tasting Corpora

We examined the importance of ideophones in tasting descriptions by comparing the frequency of the top 10 ideophones in each corpus with the frequency of the same words in the BCCWJ (The Balanced Corpus of Contemporary Written Japanese). The BCCWJ is one of the largest written corpora of Japanese, containing 104.30 million words that cover many genres [11].

Tables 2 and 3 show the frequency and adjusted frequency (per 100,000 words) of the top 10 ideophones for each corpus, as well as the ratio between adjusted frequencies for the tasting corpora and the BCCWJ. This ratio is given in the last column of the tables.

	Words (in standard romaji)	Wine Corpus		BCCWJ		Ratio
		Freq.	/100000wds	Freq.	/100000wds	
1	<i>siQkari (sikkari)</i>	616	3060.26	333	0.31	9871.8
2	<i>hoNnori (hon'nori)</i>	120	596.15	7921	7.26	82.1
3	<i>taQpuri (tappuri)</i>	77	382.53	3345	3.07	124.6
4	<i>siQtori (sittori)</i>	62	308.01	497	0.46	669.6
5	<i>huNwari (hun'wari)</i>	50	248.40	336	0.31	801.3
6	<i>simi-jimi (simi-jimi)</i>	38	188.78	514	0.47	401.7
7	<i>kiQchiri (kit'chiri)</i>	37	183.81	812	0.74	248.4
8	<i>jiQkuri (jikkuri)</i>	37	183.81	1468	1.35	136.2
9	<i>kitiN(-to) (kichin-to)</i>	29	144.07	6200	5.68	25.4
10	<i>zuQ(-to) (zutto)</i>	21	104.33	11732	10.75	9.7

Note. The capital letter 'Q' indicates adouble consonant ('ク' in Japanese). Likewise, 'N' indicates aclosed syllable ('ク' /n/, /ŋ/ in Japanese).

Freq. =observed frequency of the words in each corpus; /100,000wds = adjusted frequency per 100,000 words. The adjusted frequency of the BCCWJ is calculatedas follows: *observed frequency in BCCWJ / 2,434,619 (tokens of BCCWJ) *100,000*. Ratio =ratio between the adjusted frequencies of a word in the Wine or Sake Corpus and in the BCCWJ.

Table 2. Frequencies of the top 10 ideophones for the Wine Corpus

Name and email of the author: Hiroki Fukushima (fh@keio.jp)

¹*Real Wine Guide*, RWG, vol.19–26, 2007–2009.

Hasegawa, Ko'ichi. (2015). *Nihonshutecho* [Sake handbook], Gakken Publishing.

Kimishima, Satoshi. (2011). *Nihonshukanzen guide* [Perfect sake guide], Ikeda.

Matsuzaki, Haruo. (2000). *Nihonshu Guide Book* [Sake Guide Book], Shibata.

SSI. (2010). *Nihonshutecho* [Sake handbook], Tokyo Shoseki.

Yamamoto, Yoko. (2014). *Gensennihonshutecho* [Selected sake handbook], Sekai Bunka-sya.

Dancyu, Vol. 2, 3 (2015), Vol. 3 (2016), President.

Pen. (2015). Vol. 11, CCC Media House.

	Words (in standard romaji)	Sake Corpus		BCCWJ		Ratio
		Freq.	/100000wds	Freq.	/100000wds	
1	<i>hoNnori (hon'nori)</i>	87	173.49	7921	7.26	23.90
2	<i>siQkari (sikkari)</i>	56	111.67	333	0.31	360.20
3	<i>kiriQ(-to) (kiritto)</i>	41	81.76	56	0.05	1635.20
4	<i>torori (torori)</i>	37	73.78	158	0.14	527.00
5	<i>aQsari (assari)</i>	27	53.84	1282	1.18	45.60
6	<i>suQkiri (sukkiri)</i>	24	47.86	969	0.89	53.80
7	<i>yuQtari (yuttari)</i>	24	47.86	969	0.89	53.80
8	<i>guQ(-to) (gutto)</i>	22	43.87	975	0.89	49.30
9	<i>sarari (sarari)</i>	21	41.88	215	0.20	209.40
10	<i>saQpari (sappari)</i>	20	39.88	1421	1.30	30.70

Table 3. Frequencies of the top 10 ideophones for the Sake Corpus

As shown in Tables 2 and 3, all of the top 10 ideophones appear more frequently in tasting corpora than in the BCCWJ. The average Wine Corpus/BCCWJ frequency ratio is 1237.08, and the average Sake Corpus/BCCWJ frequency ratio is 298.89. This result is consistent with the assertions of Asano & Watanabe [1], who suggested that ideophones are more commonly used for expressing the senses of touch, olfaction, and taste.

However, this result leads to a question of why such a high ratio of ideophones exists in tasting descriptions. Ideophones are not listed as general ‘Tasting Words’ in the *Sommelier’s Textbook*, even in Japanese [10], because they can be rephrased into other adverbs. The following section attempts to explain this tendency in greater detail.

2. Co-occurrence of Ideophones in Tasting Descriptions

2.1 Method for Revealing the Roles of Ideophones Within Co-occurrence Relationships

A limited number of words are available to describe taste directly in Japanese. To make up for this lack of direct expressions, metaphorical expressions or ideophones are commonly used instead [1][2][14]. In this case, ideophones might be considered as linguistic strategies. However, the details of these strategies, or the roles of ideophones, are not immediately clear.

Therefore, to determine the roles of ideophones, we identified words that tended to co-occur in sentences containing ideophones (hereinafter called ‘ideophone-sentences’). We defined significantly frequent words in ideophone-sentences as words that needed to be modified by ideophones. The role of ideophones can be defined as describing the detailed nuances of those words. Note that the data and tables in this section are taken from our previous study in Japanese [6].

2.2 Target Corpora

As shown in Fig. 1 and Table 9, three different corpora were used: (a) the WC (Wine Corpus) represents the full data set; (b) the WIPC (Wine Ideophone Paragraph Corpus) is constructed from paragraphs containing ideophones extracted from the WC; and (c) the WISC (Wine Ideophone Sentence Corpus) is constructed from ideophone-sentences extracted from the WIPC.

In the same way, the SISC (Sake Ideophone Sentence Corpus) was constructed based on 9,799 tokens. However, an SIPC was not constructed because (a) the average number of sentences per paragraph in the Sake Corpus was relatively small (1.93 sentences/paragraph in the SC; 5.84 sentences/paragraph in the WC), such that the gaps in size and content between the SIPC and SISC would be too small, and (b) the size gap between the SC and SISC was not large enough to examine statistically.

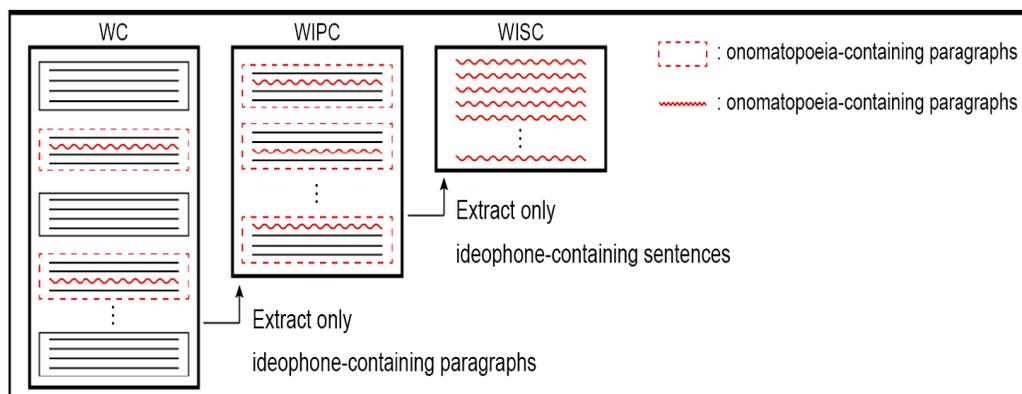


Figure 1. The Composition and Making Process of WC, WIPC and WISC

	WC	WIPC	WISC	SC	SISC
Tokens	201,294	100,977	27,928	50,147	9799
Types	9,449	6,440	2,778	3,047	1,042
Sentences	11,421	5,875	1,373	2,552	434
Paragraphs (Brands of Wine)	1,957	986	—	1,167	—
Average Frequency	10.50 (64.55)	7.91 (41.19)	5.38 (24.33)	8.47 (34.92)	5.00 (15.63)

Note. ‘Paragraphs’ corresponds to the number of different wine brand descriptions.

Standard Deviation in parentheses. Translated and reprinted from [6]

Table 4. Detailed Data of Wine and Sake Corpora

2.3 Analysis

2.3.1 Analysis method

Using chi-square tests (with Yates’ correction), we examined whether significant differences in word frequency existed between the corpora (first, WIPC vs. WISC; second, SC vs. SISC). The top 150 words of each corpus were examined, excluding the following types of words: (a) duplicated words between corpora, (b) postpositional particles, (c) ending parts of compound verbs, and (d) ideophones (because it is self-evident that ideophones are contained in sentences that contain ideophones). Because of this duplication, many of the 300 words (150 of each corpus) were eliminated. The final number of analysed words was 157 in the WISC-WISC and 180 in the SC-SISC.

2.3.2 Notes for the Analysis of the Wine Corpus

Compared to the number of tokens in the WC (more than 200,000), the number in the WISC (approximately 28,000) is considerably smaller. A large difference in corpus size can cause incorrect results in the chi-square tests. Thus, we analysed the difference between the WIPC (100,000 types) and the WISC. The WIPC is a corpus of ideophone-containing paragraphs, which means that the co-occurrence frequency was tested under more strict conditions.

2.3.3 Notes for the Analysis of the Sake Corpus

Some of the words in the top 150 words of the SISC occurred five or fewer times. Because chi-square tests require more than five samples for exact results, Fisher’s exact tests were performed to analyse low-frequency words instead. Relative risks (*RR*) and odds ratios (*OR*) were reported for the Fisher’s tests.

2.4 Results

2.4.1 Wine Corpus

Chi-square analyses on the top 150 words of the WISC and WIPC revealed significant differences between the corpora for 39 words. Of these 39 words, 20 were high-frequency words in the WISC, and 19 were high-frequency words in the WIPC ($p < .05$) (Tables 5 and 6).

Significantly high-frequency words in the WISC include the 25 words shown in Table 5. Significantly high-frequency words in the WIPC include the 19 words shown in Table 6.

	frequency observed value		adjusted frequency (per 10,000 Words)		results of tests between WISC-WIPC	
	WIPC	WISC	WIPC	WISC	χ^2 stats	<i>ps</i>
aroma (houkou)	20	12	1.98	4.30	3.84	<.050
mild (odayaka)	30	16	2.97	5.73	3.92	.048
nutrient (yohbun)	17	11	1.68	3.94	4.14	.042
grip (grip)	30	17	2.97	6.09	5.00	.025
mineral (mineral)	430	149	42.58	53.35	5.43	.020
firm (tojiru)	100	13	9.90	4.65	6.29	.012
acid (san)	286	106	28.32	37.95	6.38	.012
complexity (fukuzatsu-sei)	100	45	9.90	16.11	6.97	.008
luxuries (koukyuu-kan)	56	29	5.55	10.38	7.05	.008
sweetness (amami)	77	38	7.63	13.61	8.12	.004
construction (kouzou)	31	20	3.07	7.16	8.25	.004
dark color fruit (kuro-aka-kajitsu)	68	35	6.73	12.53	8.50	.004
easy (yasashii)	286	111	28.32	39.75	8.93	.003
umami (umami)	106	53	10.50	18.98	12.09	<.001
flavor (fuumi)	226	96	22.38	34.37	12.15	<.001
fruit (kajitsu)	342	141	33.87	50.49	15.74	<.001
structure (kokkaku)	45	30	4.46	10.74	13.80	<.001
umami (uma-mi)	61	36	6.04	12.89	12.75	<.001
smell [verb] (kaoru)	138	70	13.67	25.06	16.94	<.001
mature (noru)	64	40	6.34	14.32	16.32	<.001
center (man-naka)	57	41	5.64	14.68	22.34	<.001
terroir (daichi)	214	107	21.19	38.31	25.13	<.001
flavor of terroir (daichi-ka)	201	108	19.91	38.67	31.44	<.001
taste (aji)	650	245	64.37	87.73	16.97	<.001
flavor (ajiwai)	444	171	43.97	61.23	13.36	<.001

Note. $p < .05$, degrees of freedom= 1. Translated and reprinted from [6]

Table 5. Significantly High-frequency Words in WISC

2.4.2 Sake Corpus

Chi-square analyses and Fisher's exact tests on the top 150 words of the SC and SISC revealed significant differences between the corpora for 21 words. Of these 21 words, 16 were high-frequency words in the SISC, and five were high-frequency words in the SC ($p < .05$) (Tables 7 and 8).

High-frequency words in the SISC include the 16 words shown in Table 7). High-frequency words in the SC include the five words shown in Table 8.

	frequency (observed value)		adjusted frequency (/10,000 Words)		results of tests between WIPC -WISC			
	WIPC	WISC	WIPC	WISC	χ^2 -stats	<i>ps</i>	<i>RR</i>	<i>OR</i>
slightly (sukoshi)	239	48	23.67	17.19	3.85	>.050		
Bourgogne (Bourgogne)	88	13	8.71	4.65	4.10	.043		
faintly (jakkān)	46	5	4.56	1.79	-	.041	1.2	2.5
compare (kuraberu)	60	7	5.94	2.51	4.33	.037		
brew (tsukuru)	99	15	9.80	5.37	4.38	.036		
tasting (shi'in)	80	10	7.92	3.58	5.31	.021		
great (migoto)	175	30	17.33	10.74	5.57	.018		
cask tasting (kareta)	42	3	4.16	1.07	-	.011	1.2	3.9
really (hontou-ni)	59	5	5.84	1.79	-	.006	1.2	3.3
high (takai)	161	24	15.94	8.59	7.74	.005		
now (ima)	160	23	15.85	8.24	8.41	.004		
excellent (subarashii)	98	10	9.71	3.58	9.08	.003		
quality (hinshitsu)	118	12	11.69	4.30	11.13	.001		
brand (meigara)	96	8	9.51	2.86	11.16	.001		
vineyard (hatake)	57	3	5.64	1.07	-	.001	1.2	5.3
price (kakaku)	105	6	10.40	2.15	16.36	<.001		
wine (wine)	733	138	72.59	49.41	17.17	<.001		
be made (dekiru)	182	14	18.02	5.01	23.54	<.001		
year (nen)	420	56	41.59	20.05	27.01	<.001		

Note. Underlined words appeared five or fewer times. $p < .05$, degrees of freedom = 1.
Translated and reprinted from [6]

Table 6. High-frequency words in WISC

3. Discussion of the Co-occurrence of Ideophones

3.1 Wine Corpus

In general, the WISC seems to contain more words that express tastes or flavours directly than does the WIPC. However, the WIPC contains more words for brewing or grades than does the WISC.

3.1.1 High-frequency words in WISC

In Table 5, firm (*tojiru*), construction (*kouzou*), structure (*kokkaku*), and centre (*man-naka*) were not considered to be tasting

expressions. However, these words do represent different kinds of flavour or taste in Japanese. For example, the word ‘firm’ means a stronger sense of tannins, while the words ‘construction’ and ‘structure’ stand for the complex relationships of elements such as acid, tannin, glycerin, and body that create the texture and mouthfeel of wine. Finally, ‘centre’ is used for the movement or position of wine on the tongue.

	frequency (observed value)		adjusted frequency (per 10,000 Words)		results of tests between SC and SISC		
	SC	SISC	SC	SISC	<i>ps</i>	<i>RR</i>	<i>OR</i>
hard (katai)	8	5	1.60	5.10	.048	0.74	0.31
after flavor (atokuchi)	149	42	29.71	42.86	.039	0.93	0.69
pierce (tsuranuku)	7	5	1.40	5.10	.034	0.70	0.27
uprise (tachi-agaru)	18	9	3.59	9.18	.032	0.80	0.39
smart (keimyou)	22	10	4.39	10.21	.031	0.82	0.43
sharp-finishing (kire-aji)	57	20	11.37	20.41	.030	0.89	0.56
lasting (nokoru)	20	7	3.99	7.14	.030	0.89	0.56
finishing (hiki)	30	13	5.98	13.27	.021	0.83	0.45
clear-finishing (isagiyoi)	11	7	2.19	7.14	.019	0.73	0.31
flavor (kouki)	206	58	41.08	59.19	.016	0.93	0.69
dry (karai)	63	24	12.56	24.49	.008	0.87	0.51
odd taste (zatsumi)	26	13	5.18	13.27	.008	0.08	0.39
arise (tatsu)	142	45	28.32	45.92	.007	0.91	0.62
drink (nomu)	181	56	36.09	57.15	.004	0.91	0.63
center core (shin)	14	10	2.79	10.21	.003	0.70	0.27
mouth (kuchi)	141	48	28.12	48.98	.002	0.89	0.57

Note. $p < .05$, degrees of freedom = 1. Translated and reprinted from [6]

Table 7. High-frequency words in SISC

From the WISC, we made the following observations: (a) all of the high-frequency words in the WISC are expressions about the taste or flavour of wine, rather than brewing or grades of wine; (b) the ideophones co-occur with relatively abstract expressions of taste. Tasting words for wine seem to contain various detailed concrete words (ex. *lemon*, *cinnamon*, and so on). However, as seen in Table 5, no such concrete elements occur in relation to words for grip, structure, fruit, and other aspects, which appear to be expressions of taste but are rather abstract expressions among wine tasting words.

3.1.2 High-frequency words in WIPC

The words shown in Table 6 are construed as low co-occurrence words with ideophones. In comparison with those in the WISC, the high-frequency words in the WIPC contain very few words for taste or flavour (the only exceptions being ‘kareta’ or ‘cask taste’).

In contrast, Table 6 does contain several words for brewing and standards or grades of wine. While ‘Bourgogne (*Bourgogne*)’ is a famous wine production region, the following words reference brewing or grades of wine: brew (*tsukuru*), quality (*hin-shitsu*), brand (*meigara*), vineyard (*hatake*), price (*kakaku*), wine (*wine*), be made (*dekiru*), and year (*nen*). The details of these words can be described using numbers or fixed expressions. The vivid nuance expressed using ideophones is unnecessary.

We also observed a high frequency of adverbial expressions (Table 6). Six adverbs were identified: slightly (*sukoshi*), faintly

(*jakkan*), great (*migoto*), really (*hontou-ni*), high (*takai*), and excellent (*subarashii*). This tendency might be related to the adverbial use of ideophones. The duplication of the same function is therefore not needed in a single sentence.

	frequency (observed value)		adjusted frequency (per 10,000 Words)		results of tests between SC and SISC		
	SC	SISC	SC	SISC	<i>ps</i>	<i>RR</i>	<i>OR</i>
dying (<i>kareru</i>)	25	0	4.99	0.00	.025	1.20	-
cool[verb] (<i>hiyasu</i>)	27	0	5.38	0.00	.016	1.20	-
match (<i>au</i>)	26	0	5.18	0.00	.015	1.20	-
use (<i>shiyou</i>)	29	0	5.78	0.00	.010	1.20	-
warming up (<i>kan</i>)	60	3	11.96	3.06	.010	1.14	3.91

Note. $p < .05$, degrees of freedom = 1. Translated and reprinted from [6]

Table 8. High-frequency words in SC

3.2 Sake Corpus

The results from the Sake Corpus further indicate that ideophones co-occur with words of taste and flavour, which are considered to be relatively abstract expressions.

The primary concern when commenting on sake is its taste or flavour. Therefore, mentions of grades or vineyard locations do not appear in the Sake Corpus as much as in the Wine Corpus. Sake tasting descriptions thus provide clearer insight into the roles of ideophones. In the Sake Corpus, as in the Wine Corpus, strong connections occur between the co-occurrence of ideophones and words for taste. All 16 words in Table 7 describe tastes. Alternatively, four of the five words in Table 8 do not describe tastes or flavours.

The tests in the Sake Corpus show that ideophones are commonly used to express the sense of a flavour's 'appearance' or 'finish'. Typical 'appearing' taste descriptions use words such as rising up (*tachi-agaru*) and arise (*tatsu*), while 'finishing' taste descriptions include words such as after flavour (*atokuchi*), sharp-finishing (*kire-aji*), remain (*nokoru*), finishing (*hiki*), and clear-finishing (*isagiyo*). We can therefore infer that ideophones are used for describing both the changing of tastes and the tastes themselves.

In summary, we draw the following conclusions about the collocation of ideophones:

- Ideophones are used to modify words for taste, rather than words for brewing or grades.
- Ideophones co-occur with relatively abstract expressions of taste.
- Ideophones are used to express the sense of 'appearing' or 'finishing' of flavours (specifically in the Sake Corpus).

We conclude that ideophones are employed to describe the changing of tastes rather than their stable states. This point distinguishes ideophones from general adverbs, which describe mainly the states and manners of objects or events.

4. The Morphology of Ideophones in Tasting Descriptions

4.1 The Classifications of Japanese Ideophones

Japanese ideophones can be classified into three main categories: phonomimetic (*gi-on-go* in Japanese), phenomimetic (*gi-tai-go*), and psychomimetic (*gi-joh-go*). In actuality, the borders of these categories are often ambiguous; one ideophone can be used in two or three categories. It is thus nearly impossible to investigate the real composition rate of these categories automatically. Huan [9] determined the compositions of these categories based on relatively limited linguistic data by manually counting up the ideophones seen in four Japanese novels. In this section, we contrasted the compositions of ideophones in tasting descriptions with that in Huan's Novel Corpus (Table 5). Note that, in Huan's Novel Corpus, the frequency of phonomimetic ideophones is 84, that of phenomimetic ideophones is 1,288, and the psychomimetic classification is not used. (Psychomimetic usage is integrated into phenomimetic usage.)

	Novel Corpus		Wine Corpus		Sake Corpus	
	Type	Freq.	Type	Freq.	Type	Freq.
Phono-mime	-	84	4	2	1	1
Pheno-mime	-	1,288	157	1,575	50	543
Psycho-mime	-	-	6	61	0	0

|-----| * (a)

|-----| * (b)

Note. * = $p < .05$

Table 9. Classifications of ideophones in the Novel Corpus, Wine Corpus, and Sake Corpus

4.1.1 Methods and Results

Using Fisher's exact tests, we examined whether significant differences in the frequency of each classification existed between the Novel Corpus and both the Wine Corpus and Sake Corpus.

Fisher's exact test on the classifications revealed the following frequency differences for corpus types: (a) fewer phonomimetic ideophones in the Wine Corpus than in the Novel Corpus ($OR = 53.348, p < .0001$) and (b) fewer phonomimetic ideophones in the Sake Corpus than in the Novel Corpus, ($OR = 0.613, p < .0049$). In contrast, phonomimetic ideophones appear more frequently in the tasting corpora. Note that OR indicates the odds ratio and CI indicates the confidence interval.

4.1.2 Discussions on the Classifications of Japanese Ideophones

One primary explanation for the result that fewer phonomimetic and more phenomimetic ideophones are present in the tasting corpora is that wines and sakes are liquids. Beverages give almost no real sound in the mouth. However, ideophones are not less vivid when expressing beverage taste. Instead of representing the real sound of the environment, phonomimetic and psychomimetic ideophones express the texture, manner of change, and impressions of tasting. Based on this framework, the next section examines the detailed functions and roles of ideophones.

4.2 The Word Bases of Ideophones

4.2.1 Methods

Japanese ideophones can be further classified into three types of word base: A-Type, AB-Type, and AR-Type [8]. These types are determined by the word's number of syllables and consonants. The A-Type contains ideophones with a single consonant and only one or two syllables. The AB-Type contains ideophones with two or more different consonants and two or three syllables. The AR-Type is a variation of the AB-Type in which the consonant of the second syllable is 'r'.³ Table 10 shows

Base Type	BCCWJ (2009)	Wine Corpus	Sake Corpus
A-type	289 (18.43%)	29 (17.57%)	9 (18.00%)
AB-type	928 (59.18%)	106 (64.24%)	33 (66.00%)
AR-type	317 (20.22%)	30 (18.18%)	8 (16.00%)
others	34 (2.17%)	-	-
Total	1,568 (100%)	165 (100%)	50 (100%)

Table 10. Base Type Compositions of the BCCWJ, the Wine Corpus, and the Sake Corpus

the base type compositions in the BCCWJ, Wine Corpus, and Sake Corpus. The base type classification of the BCCWJ (version 2009) is based on Huan's study [9].

4.2.2 Results

Using one-way chi-square tests, we examined whether significant differences existed in the ratios of base type composition between the BCCWJ and both the Wine Corpus and Sake Corpus. Chi-square tests revealed no significant difference either between the BCCWJ and the Wine Corpus (A-Type: $\chi^2=0.03$, $p=.870$; AB-Type: $\chi^2=1.38$, $p=.239$; AR-Type: $\chi^2=0.27$, $p=.604$), or between the BCCWJ and the Sake Corpus (A-Type: $\chi^2=0.00$, $p=.976$; AB-Type: $\chi^2=0.42$, $p=.519$; AR-Type: $\chi^2=0.38$, $p=.537$). The results of these tests suggest that the numbers of syllables and consonants themselves do not alone account for the characteristics of ideophones intasting descriptions.

4.3 The Patterns of Ideophones

We further analysed the patterns of ideophones using Fisher's exact tests. Some patterns exist for each ideophone base type. For example, A-Type ideophones are classified into six patterns; *CVQ*, *CVN*, *CVRN*, *CVRQ*, and other. In this case, 'CV' refers to a single consonant. 'N' is a closed syllable, 'ㄥ' (/n/) in Japanese. Q is a double consonant, "ㄷ" in Japanese. R is a prolonged sound, 'ー' in Japanese. Table 11 shows the compositions of patterns foreach base type.

4.3.1 Method and Results

To determine whether characteristic patterns existed in the tasting descriptions, we performed Fisher's exact tests on each pattern. The tests were performed between the BCCWJ and both the Wine Corpus and Sake Corpus. Each test was performed independently. Tests were not performed on the A-Type ideophones of the Sake Corpus due to their limited number.

Patterns	BCCWJ (2009)	Wine Corpus	Sake Corpus	ps in BCCWJ-WC
<i>CVQ</i>	18 (46.15%)	10 (34.48%)	9(100%)	.614
<i>CVN</i>	9 (23.08%)	8 (27.59%)	0	1.000
<i>CVRN</i>	7 (17.95%)	6 (20.69%)	0	.754
<i>CVRQ</i>	0	2 (6.90%)	0	.156
others	1 (2.56%)	2 (6.90%)	0	.559
TOTAL	4 (10.26%)	1 (3.45%)	0	.144
TOTAL	39	29	9	-

Note. Numbers represent the types (not tokens) of ideophones for each pattern.

Table 11. Pattern Compositions of A-Type Ideophones

For A-Type ideophones, Fisher's exact tests found no significant differences in the *CVQ* patterns. This result is likely due to the fact that all of the ideophones in the Sake Corpus followed the *CVQ* pattern. This pattern may be characteristic, but more data are needed to be certain.

Further comparisons for the AB-Type did reveal several statistically significant differences. The *CV CV' - CV CV'* pattern appeared more frequently in the BCCWJ than in either the Wine Corpus ($OR = 1.929$, $95\% CI = 1.055-1.715$; $p = .027$) or the Sake Corpus ($OR = 8.571$, $95\% CI = 1.200-1.630$; $p < .001$). In contrast, the *CV CV' Q* pattern was used more frequently in the Wine Corpus than in the BCCWJ ($OR = 0.163$, $95\% CI = 0.085-1.080$; $p = .015$).

Similar results were also found for the AR Type. The *CV rV - CV rV* pattern appeared more frequently in the BCCWJ than in either the Wine Corpus ($OR = 8.196$, $95\% CI = 1.164-8.928$; $p = .002$) or the Sake Corpus ($OR = 21.111$, $95\% CI = 1.006-6.071$; $p = .001$). The characteristic patterns in the tasting corpora were the *CV rV Q* pattern in the Wine Corpus ($p = .002$) and the *CV rV ri* pattern in the Sake Corpus ($OR = 0.079$, $95\% CI = 0.200-1.121$; $p < .001$).

¹ Examples of AR Type: *KaRiQ-to* (crispy), *SaRa-SaRa* (smooth feelings), etc

4.3.2 Discussion on Types of Ideophones

Among the threetypes and 17 patterns of ideophones, two patterns (*CV CV' – CV CV'* and *CV rV – CV rV*) were not characteristic in the tasting corpora. These two patterns shared the same structure, employing a repetition of sounds. ¹In contrast, the *CV CV' Q* and *CV rVQ* patterns were characteristic in the Wine Corpus. These patterns also share a similar sound structure, containing a double consonant.

Patterns	BCCWJ (2009)	Wine Corpus	Sake Corpus	<i>ps in BCCWJ-WC</i>	<i>ps in BCCWJ-SC</i>
<i>CV Q CV' ri</i>	43 (36.75%)	39 (38.46%)	16 (48.48%)	.889	.233
<i>CV CV' – CV CV'</i>	54 (46.15%)	30 (30.77%)	3 (9.09%)	.027*	< .001*
<i>CV Q CV' ra</i>	2 (1.71%)	2 (1.92%)	2 (6.06%)	> .999	.210
<i>CV CV' N</i>	4 (3.42%)	2 (1.92%)	1 (3.03%)	.687	> .999
<i>CV CV' Q</i>	2 (1.71%)	11 (9.62%)	2 (6.06%)	.015*	.210
<i>CV N CV' ri</i>	9 (7.69%)	13(11.54%)	4(12.12%)	.365	.484
<i>CV CV' ri</i>	3 (2.56%)	5 (3.85%)	3(9.09%)	.709	.121
<i>CV CV'</i>	0	2 (0.96%)	1 (3.03%)	.471	.220
<i>CV C'V ri – CV C'V ri</i>	0	1 (0.96%)	1 (3.03%)	.471	.220
TOTAL	117	106	33	-	-

Note. * = $p < .05$.

Table 12. Pattern compositions of AB-Type Ideophones

Patterns	BCCWJ (2009)	Wine Corpus	Sake Corpus	<i>ps in BCCWJ-WC</i>	<i>ps in BCCWJ-SC</i>
<i>CV rV – CV rV</i>	38 (92.68%)	19 (60.71%)	3 (37.50%)	.002*	.001*
<i>CV rV Q</i>	0	4 (17.86%)	1 (12.50%)	< .001*	.163
<i>CV rV ri</i>	3 (7.31%)	7 (21.43%)	4 (50.00%)	.144	< .001*
TOTAL	41	30	8		

Note. rV: the mora of “ra”, “ri”, “ru”, “re” and “ro”. * = $p < .01$.

Table13. Pattern compositions of AR-Type Ideophones

Typical repetition types include words such as ‘*kuru-kuru*’ (continuous turning motion; rolling or spinning around). Generally, repetition types are thought to be motivated by the continuity or repetition of real, physical sound, movements, or situations [12]. The lower frequency of repetition in the tasting corpora suggests that ideophones are not used for representing the continuous states or conditions of tastes.

In contrast, typical double consonant typesinclude words such as ‘*kuruQ*’. The double consonant in ideophones can represent the momentary breaking point or turning point of sounds, movements, or situations[12]. For example, while *kuru-kuru*reflects a continuing rolling motion, *kuruQ* evokes a single turn of an object. The high frequency of double consonants types thus suggests that ideophones can be usedto describe or point to the turning point or breaking point of the temporal flow of tastes.

Based on thesetypes, we find the following general tendency of ideophones in tasting descriptions: ideophones tend to be used to represent the breaking or turning points of taste, rather than continuous states or conditions.

5. Conclusion

The purpose of this study was to examine the morphological features and co-occurrence tendencies of Japanese ideophones in relation to descriptions of taste. Morphology is one of the most fundamental themes in studies of ideophones, but little previous work has examined the morphology of ideophones in taste descriptions.

In this study, using the Sake Corpus and Wine Corpus, we determined the morphological features and co-occurrence tendencies of Japanese ideophones in descriptions of taste. We found the following co-occurrence tendencies for Japanese ideophones in the tasting corpora[6]:

- Ideophones are used to modify words for taste, rather than words for brewing or grades of wines and sakes.
- Ideophones co-occur with relatively abstract expressions of taste.
- Ideophones are used to express the sense of ‘appearing’ or ‘finishing’ of flavours (specifically in the Sake Corpus).

These results suggest that ideophones metaphorically bridge the taste and sound domains to represent comparatively low-salience elements of tasting phenomena, such as the changing process (e.g., appearing and finishing, not the taste itself).

Based on these findings, we investigated the morphological features of ideophones in the tasting descriptions. Morphological analysis is the most fundamental approach for studying the iconicity of ideophones. We found the following results:

- A high frequency of ideophones is observed in the tasting corpora. Ideophones seem to play an important role in tasting descriptions.
- Significantly fewer phonomimetic ideophones appear in the tasting corpora than do phenomimetic ideophones. These ideophones are thought to represent the manner of the taste and not the real sound in the mouth.

These two points may seem conflicting: ideophones are highly used in tasting descriptions but do not represent sounds themselves. What is, then, represented by these ideophones?

Ideophones are not only motivated by real sounds. Therefore, we investigated the forms (base types and patterns) of the ideophones. We found the following formal features of ideophones in this study:

- No significant differences exist between the base types of ideophones.
- Several features characterise the patterns of ideophones: (a) the repetition types ($CV CV' - CV CV'$ and $CV rV - CV rV$) appear more frequently in the BCCWJ, and (b) the double consonant patterns ($CV CV' Q$ and $CV rV Q$) are characteristic in the tasting corpora. These ideophones tend to represent the breaking points or turning points of taste rather than continuous states or conditions.

References

- [1] Asano, M., & Watanabe, J. (2014). Chikaku to gengo (perception and language). In M. Imai, & N. Saji (Eds.), *Gengo toshintaisei (Language and Embodiment)* (pp. 63-92)
- [2] Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition*, 130(2), 266-270.
- [3] Atoda, T., & Hoshino, K. (1993). *Giongo gitaigo tsukaikata jiten (The usage guide to Japanese onomatopoeias)*. Sohtaku.
- [4] Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, 6(10), 654-672.
- [5] Dingemanse, M., & Akita, K. (2016). An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics*, doi:10.1017/S002222671600030X

⁴The repetition patterns of the Wine Corpus include *shimi-jimi*, *don-don*, *pun-pun*, *fuwa-fuwa*, etc. The repetition patterns of the Sake Corpus include *sui-sui*, *chira-chira*, etc.

- [6] Fukushima, H., Mutsumi, I., & Shigenori, T. (2016). The roles of sound symbolisms in the tasting descriptions. *Transactions of the Japanese Society for Artificial Intelligence*, 30 (6).
- [7] Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, CA: CSLI Publications.
- [8] Hira, M. (2004). Gendaigo ni okeru giongo no pattern ni tsuite (patterns of onomatopoeia in modern Japanese). *Bulletin of the Center for Japanese Language, Doshisha University*, 4(12), 17-29.
- [9] Huan, H. (2011). Consideration about the basic vocabulary of the onomatopoeia: A study using balanced corpus of contemporary written Japanese (monitor published version). *Japanese Studies: Research and Education Annual Report*, 15, 17-39.
- [10] Japan Sommierier Association. (2015). *Japan sommierier association textbook 2015* Japan Sommierier Association.
- [11] Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., . . . Den, Y. (2014). Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation*, 48(2), 345-371.
- [12] Ono M. (Ed.) (2007). *Nihongo onomatope jiten (the dictionary of Japanese onomatopoeias)*, Shogakukan.
- [13] Ohsawa, Y., Kido, H., Hayashi, T., & Liu, C. (2013). Data jackets for synthesizing values in the market of data. *Procedia Computer Science*, 22, 709-716.
- [14] Seto, K. (2003). *Kotoba wa aji wo koeru (words excel the taste)* Kaimeisha.
- [15] Tamori, I., & Schourup, L. (1999). *Onomatope -keitai to imi (onomatopoeia -the forms and the meanings)*. Kuroshio.
- [16] Voeltz F. K. E., Christa Kilian-Hatz (Eds.). (2001). *Ideophones*, John Benjamins.