

Available online at www.sciencedirect.com**ScienceDirect**

Procedia Computer Science 96 (2016) 1067 – 1074

Procedia
Computer Science

20th International Conference on Knowledge Based and Intelligent Information and Engineering Systems

A Framework to Collect Japanese Expression for Food Taste and Texture

Naoki SHINO^a, Ryosuke YAMANISHI^{b,*}, Yoko NISHIHARA^b, Junichi FUKUMOTO^b^aGraduate school of Information Science and Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu 525-8577, Japan^bCollage of Information Science and Engineering, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu 525-8577, Japan

Abstract

Japanese has more variety of expressions for food; the expressions have been changed through the times. In Japanese, “metaphor,” “onomatopoeia,” and “adjectives generally not for food” are often used to express the taste and texture of food. Which and how the expression is used depends on individual sensibility based on location and generation; especially, the youngsters sometimes use new expressions. Since, it is difficult to extract Japanese taste and texture expressions from Web reviews based on a fixed dictionary. This paper proposes an interactive framework to collect Japanese taste and texture expressions for food. Using the proposed framework, varied and many expressions would be semi-automatically collected. Based on the collected expressions, the sentiment analysis for food will be facilitated.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of KES International

Keywords: Cooking, Recipe review, Expression for food

1. Introduction

It has been recently popular to get food information from varied websites for each objective. In order to enjoy eating, buzz marketing site is used to choose a restaurant, and recipe site is used to cook a dish. Also, Wikipedia and homepage of restaurants are used to acquire knowledge of food nutrients: calorie of a dish and nutrient composition. It has been increasing that the websites for food written in Japanese are browsed, for example Japanese restaurant and recipe for Japanese traditional cooking, i.e. *Washoku*. Behind the trend is the fact that *Washoku* has been nominated to the UNESCO heritage list (p.4) and on a world-scale trend. As the another reason, Japanese recipe websites (e.g., “Cookpad” and “Rakuten recipe”) have started to provide their recipe data with researchers; the research for food engineering becomes corpus-based research. From both sides of enjoying and research, Japanese information for food has been highly demanded.

The field of food engineering is recently more active in Japan. Such researches are not limited to just Cooking Science but also in Information Science, for example, a research community of Multimedia on Cooking and Eating

* R. Yamanishi. Tel.: +81-77-561-3416.

E-mail address: {n.shino@nlp.is, ryama@media, nisihara@fc, fukumoto@media}.ritsumei.ac.jp

Activities, Human Communication Group, IEICE. Interpreting food as one of media, the researchers discuss about media processing especially for food based on natural language processing, image processing and bioinformatics. Mori *et al* proposes a method to recognize Japanese recipe terminology; for example, cooking action and name of ingredients or cookware¹. Their system was developed by using machine learning with texts of cooking instructions. They also developed another system which translate recipe text to words, that is, a morphological analyzer for recipe document. However, their systems do not sufficiently cover expressions for tastes and textures of foods (i.e., food expression), which are not frequently occurred in cooking instructions. If their system covers food expressions, natural language understanding might be expanded for all of text describing foods from only recipe text.

The natural language understanding would contribute to a machine translation for Japanese text in food domain; the more natural language understanding is improved the more the translation is improved. Expansively, the collected food expressions will be available to rephrase language-depended expressions such as onomatopoeia. For food taste and texture, some language-depended expressions are often used in the document. Such expressions can not be directly translated but indirectly explained by using example based on relationships between food expressions and food. For example, the explanation of difference between カリカリ (*Kari-kari*) and サクサク (*Saku-saku*) can be described as follows; “Both *Kari-kari* and *Saku-saku* roughly mean crispy. *Kari-kari* is often used for nuts and well baked bacon. *Saku-saku* is often used for pork cutlet, baked baguette, and cookie.”

Then, Japanese food expressions have some problems based on characteristics and expressions variety in Japanese. The problems must be solved for information science for food. To tackle such problems, knowledge of Japanese words and phrases for food should be summarized. This paper proposes an interactive framework to collect sentences which include words and phrases for food expressions. The proposed framework enables us to semi-automatically collect food expressions with its disambiguation. It is expected that the collected food expressions will effectively contribute machine translation or natural language understanding in food domain as the learning data of machine learning for Japanese recipe recognition.

2. Problems to solve

The goal of this study is to analyze and translate texts including food expressions. Machine learning requires sample documents including correct food expressions. It is difficult to automatically determine whether a text correctly concerns taste or texture because of the following problems especially in Japanese;

- (a) There are too many words and phrases that express texture of food.
- (b) Words are connected to each other.

Japanese has too many texture words and phrases. National Food Research Institute (NFRI)¹ defines 445 Japanese words and phrases as Japanese food texture. On the other hand, English food textures have been surveyed². NFRI reports that English food texture are expressed by only 77 words. It must be difficult to recognize little differences between each Japanese food texture for non-native Japanese. Because of the variety of texture words and phrases, food texture for a same food depends on individuals even among native Japanese. Since, it is impossible to directly translate a Japanese food texture to an English one. And, Hayakawa *et al* describe that the number of food textures are increasing and decreasing³. They discovered the increase and decrease by comparing their test on 2003 and the test conducted by Yoshikawa *et al*⁴. The increasing and decreasing have relation to too many texture words and phrases. Moreover, some words and phrases have multiple meanings of food texture and other. For example, “まつたり (*Mattari*)” means soft and full-bodied in a domain of food. Alternatively, the word has a meaning of relaxed atmosphere. Whether a word/phrase is food texture or not must be distinguished; this is the difficult point to solve.

As another problem, a word is connected to another word without space or segmentation in Japanese unlike English. Therefore, Japanese text can not be translated without splitting sentences by each words. In most case of natural language processing, morphological analysis is conducted to split a sentence by each words. The morphological

¹ <https://www.naro.affrc.go.jp/english/nfri/index.html>

analyzer for specific domain should be modeled with a learning data for the specific domain. The learning data for food domain has been not sufficiently progressed yet.

From the above facts, it is difficult to collect documents for food expressions and to determine whether a text rightly expresses taste/texture without humans. So, existing food expressions databases such a NFRI database are structured based on bibliographic survey and questionnaire survey to food researchers and experts of texture research; it requires too much effort, and has been not advanced. The work would be more efficient by automatic collection of candidate documents for food expressions: humans have to only determine whether the document rightly express food expressions or not. To develop such a semi-automatic framework, automatic collection of candidate documents and criteria for food expressions representing taste and texture should be defined.

3. The proposed framework

The proposed framework considers the following aspects;

- What sentences should be automatically collected as the candidate document for food expressions?
- How does the user determine whether the candidates are correct food expressions or not?
- How does the user correct a wrong candidate for food expressions into the appropriate one?

The detail of the aspects will be shown in section 3.1 and 3.2. And the demonstration of the proposed framework will be shown in section 3.3.

3.1. Sentences for food expressions to be automatically collected

As candidate documents to be verified by humans, the framework automatically collects sentences including strings partially matched with food taste and texture expressions in the prepared dictionary. To prevent missing sentences should be emphasized than preventing to collect wrong sentences for this framework. Partial match is effective for this framework because it picks up more sentences.

For the first step of developing the framework, a dictionary including a certain number of food taste and texture must be prepared. In this paper, food textures defined by National Food Research Institute is prepared as the primal dictionary of food expressions. The food textures are collected using questionnaires by Hayakawa *et al.* The dictionary will be expand and improved by interaction with humans; the expansion will be detailed in section 3.2.

As the correct expression for food taste and texture, the proposed framework provides the following definitions;

Definition 1 The sentence must include evaluation of food that the reviewer eat.

Definition 2 The sentence must include string a string partially matched with a term concerning food taste and texture in the prepared dictionary with its right usage.

As the user accept only the sentences that meet both of the above definitions are accepted, and then the sentences that rightly express food taste and texture are collected. In the next subsection, how to handle the sentence does not meet definition 2 but meet definition 1 will be explained.

3.2. How to correct wrong candidates

If a sentence does not meet definition 2 but meet definition 1, the sentence might include not the partially matched food expression but another expression. Then, the user provides other appropriate expression. If the expression provided by the user is included in prepared dictionary, the sentence should be assumed as the sentence for the expression user newly provides. If the expression provided by the user is not included in the dictionary, the input expression becomes a candidate of new expression of food taste or texture; that is, brand new food expression is discovered.

Table 1. Three examples of recipereview using kusya-kusya

Index	Original in Japanese	Translation to English
1	くしゃくしゃで食べやすいです。	This (salad) is easy to eat, it's <i>kusya-kusya</i> .
2	冷蔵庫でくしゃくしゃになっていたキャベツを発見。	I found a <i>kusya-kusya</i> cabbage in my refrigerator.
3	新玉ねぎが甘くくしゃくしゃです。	A new onion was sweet and <i>syaku-syaku</i> .

3.3. Demonstration of the proposed framework

Demonstration of the proposed framework will be showed using “くしゃくしゃ (*kusya-kusya*)” which is food texture meaning “crumpled.”

Table 1 shows three recipe reviews using *kusya-kusya*. The review 1 is a correct example sentence. In this review, *kusya-kusya* should expresses appearance that the salad became crumpled by teeth because this review said that it is easy to eat by *kusya-kusya*. Therefore, this review is a correct as example of a sentence using *kusya-kusya* as food texture.

The review 2 is a wrong example. In this case, *kusya-kusya* must express feeling for looks or touch. So, this review should be eliminated as wrong example.

The review 3 is a wrong example for *kusya-kusya*, but it expresses food texture. This review can be divided “新 (new) 玉ねぎ (onion) が (is) 甘く (sweet) しゃくしゃく (texture meaning flesh)” truthfully. So, this review must be used as example of a sentence using “しゃくしゃく (*syaku-syaku*).” Therefore, if *syaku-syaku* is included in prepared dictionary, this sentence is used as example of a sentence using *syaku-syaku*. And, if it is not included in the dictionary, it becomes a candidate of new expression of food texture.

4. User test

The user test for the proposed framework was conducted with food expressions (i.e., food taste and texture) defined in Hayakawa *et als*' work³. The purpose of this test was to collect example sentences for 120 words and phrases. The 120 words and phrases are shown in Table 2.

In this paper, 10,921,192 recipe reviews provided by COOKPAD were used as the document source. Randomly selected five candidates of example sentence for each food expression were automatically obtained from review documents. Then, the proposed framework obtained sentences including a string partially matched with food expression in the prepared dictionary. If less than five candidates for the food texture were in the review documents, all of the candidates were obtained. As the result, 381 candidates were obtained for 85 food expressions; no sentence was obtained for 35 food expressions with partial match. The marked food expression in Table 2 shows the food expression that one or more sentences were obtained for its sample.

Students belong to the department of information science in their twenties participated to the test. The participants judged 381 candidates whether the candidate was correct example for each food expression or not. When the candidate was incorrect for the intended food expression and correct for another one, the participant provided the another food expression.

4.1. Results

It takes approximate two hours for each participant in average. Table 3 and 4 each shows the overview result for obtained sentences and the one for the target expressions, respectively. In the user test, this paper prepared two types of criterion. The number of candidates judged as correct example by more than three participants are 94 sentences for 50 food expressions. With the sentences obtained through this user test, one or more correct example sentences can be related with the 50 food expressions. Using another criterion, 147 examples were judged as correct example by more than two participants; 53 sentences were additionally accepted from the result with the above strict criterion. These 147 examples were for 66 food expressions.

Through the user test, 16 sentences were found as incorrect for the intended food expressions but correct for another; the sentences were for 11 food expressions which were submitted by more than two participants. Table 5

Table 2. 120 food expressions used in the user test. The original in Japanese and its translation into English are shown in the table. The checkmark after the translation shows that one or more sentences were obtained for the food expression.

Food expression	Translation	Food expression	Translation	Food expression	Translation
厚い	thick ✓	乾いた	dry ✓	くたっ	kutax ✓
ぐによぐによ	gunyo-gunyo ✓	ゴリッ	gorix ✓	ざらつく	coarse ✓
皮ばった	impossible to translate	ぐによっ	gunyox ✓	ころころ	koro-koro ✓
さらり	sarari ✓	塊状の	like a chunk	キシキシ	kishi-kishi ✓
ごろごろ	goro-goro ✓	ざらり	zarari ✓	かちかち	kachi-kachi ✓
ギシギシ	gishi-gishi ✓	ころっ	korox ✓	サンドイッチ状の	like a sandwich
がちがち	gachi-gachi ✓	ぎっしり	gissiri ✓	くちゃくちゃ	kucha-kucha ✓
結晶状の	like a crystal	ごろっ	gorox ✓	しけた	wet ✓
粗い	coarse	ぎとぎと	gito-gito ✓	コキコキ	gito-gito ✓
ころり	korori	泡状の	like bubbles	かちんこちん	kachin-kochin ✓
ぎとっ	gitox	くちゃっ	kuchax	ごろり	gorori ✓
泡の立つ	standed bubbles	がっしり	gassiri ✓	ぐちゃっ	guchax
固形の	solid	こわい(強い)	stiff	しこっ	shikox ✓
いがいが	iga-iga	かどばった	angular	吸湿性がある	moisture absorbent
くちゅくちゅ	kuchu-kuchu	ごわごわ	gowa-gowa ✓	球状の	like a sphere
ぐちょぐちょ	gucho-gucho ✓	こちこち	kochi-kochi ✓	ごわっ	gowax ✓
吸水性がある	water absorbent	ぐちょっ	guchox ✓	こちっ	kochix ✓
こわれやすい	easily broken	しっかり	firm ✓	うろこ状の	like a scale
強靱な	tough	くっつく	stick ✓	こちんこちん	kochin-kochin ✓
しっけた	wet ✓	液状の	like liquid	からから	kara-kara ✓
切れやすい	easy cut	くにやくにゃ	kunya-kunya	ごつごつ	gotsu-gotsu ✓
ザクザク	zaku-zaku ✓	液のしたたる	dripping juice	均一な	regular
ぐにやぐにや	gunya-gunya ✓	じっとり	jittori ✓	重い	heavy ✓
からみつく	twist around ✓	くしゃくしゃ	kusya-kusya ✓	くにやっ	kunyax ✓
ザクッ	zakux ✓	じとじと	jito-jito	かくばった	angular
からり	karari ✓	ぐしゃぐしゃ	gusya-gusya	ぐにやっ	gunyax ✓
裂けやすい	easy to split	しどっ	shitox ✓	かざかざ	kasa-kasa
くしゃっ	kusyax	くにやり	kumyari ✓	じどっ	jitox ✓
がさがさ	gasa-gasa	ガリガリ	gari-gari ✓	ぐしゃっ	gusyax
ぐにやり	gunyari ✓	細かい	fine grains	ざっくり	zakkuri ✓
しなしな	shina-shina ✓	かさつく	impossible to translate	ぐずぐず	guzu-guzu ✓
くにゆくにゆ	kunyu-kunyu ✓	ゴムのような	like a gum ✓	さらさら	sara-sara ✓
かすかす	kasu-kasu ✓	ガリッ	garix ✓	ぐにゆぐにゆ	gunyu-gunyu
さらさら	zara-zara ✓	顆粒状の	like granulars	くたくた	kuta-kuta ✓
ぐにゅっ	gunyux ✓	ゴリゴリ	gori-gori ✓	しなやか	supple
硬い	hard ✓	軽い	light ✓	くだけやすい	easy bereaking
くによくによ	kunyo-kunyo	コリッ	korix ✓	ざらっ	zarax ✓
渋い	astringent ✓	堅い	hard ✓	脂っこい	fatty ✓
固い	hard ✓	口あたりがよい	mild	油っこい	greasy ✓
口ざわりがよい	mouthfeel is good	クリーミー	creamy ✓	脂っぽい	oily ✓

shows the 11 food expressions. In the 11 food expressions, eight of them were included in food expressions stored in the prepared dictionary but three expressions were newly obtained with the interactions. Such three food expressions would be candidates of new food expressions and might be collected into food expression dictionary.

4.2. Discussions

In the following subsections, the results will be discussed in detail. We focused on the results based on four types of aspects: food expressions for which correct examples were obtained, food expressions for which wrong partial-matched examples were obtained, food expressions for which no example was obtained, and candidates of new food expressions.

4.2.1. Food expressions for which correct examples were obtained

All participants judged 94 reviews as correct examples for food expressions. The 94 reviews were surely correct examples but the coverage of food expressions was less: the reviews were for only 50 out of 85 food expressions.

With the lax criterion, 147 reviews could be assumed as correct examples. The 147 reviews were for 66 food expressions, more expressions could have its usage example though they might include unsure examples. For example, “表面が乾いたぐらいでいいかもです (Maybe it is enough if the surface is dry)” is a review of a recipe for baked cookie. In this case, the object of “dry” may be not “baked cookie” but “pre-baked cookie,” that is, this review would

Table 3. The overview result for obtained sentences.

The Criterion	Number of the sentences	Examples of recipe reviews
Three participants judged as correct	94	ちょっと乾いた感じだけどおいしー (It's good taste but it is dry.) エリンギのくにくくにゆした食感がクリームにめちゃ合います! (<i>Kunyu-kunyu</i> texture of Eringi is very good match with its cream.) ガリガリとした噛みごたえが自分好みのクッキーでした (The cookie with <i>gari-gari</i> is suited with my sense.)
Two participants judged as correct	53	表面が乾いたぐらいでいいかもです (The surface is dried, then it may be good.)
		しっけたときどかに良いですね (It is suitable when it becomes damp.) 焼き過ぎてがちがちになってしまいました (It became <i>gachi-gachi</i> because of overcook.)
More than two participants judged as correct	147	

Table 4. The overview result for the target expressions. The result can be divided into four patterns: food expression for which correct examples were obtained, food expression for which wrong partial-matched examples were obtained, food expression for which no example was obtained, and new food expression.

	Number of the expressions	Examples of the expressions
Food expression for which correct examples were obtained	66	くにくくにゆ(<i>kunyu-kunyu</i>), さらさら (<i>sara-sara</i>), 硬い (<i>hard</i>), 軽い (<i>light</i>), 油っこい (<i>oily</i>), クリーミー (<i>creamy</i>)
Food expression for which wrong partial-matched examples were obtained	19	ごろごろ (<i>goro-goro</i>), くちやくちや(<i>kucha-kucha</i>), ころっ(<i>korox</i>), しなやか (<i>supple</i>), 均一な (<i>regular</i>)
Food expression for which no example was obtained	35	ころころ (<i>koro-koro</i>), いがいが ^s (<i>iga-iga</i>), 粗い (<i>coarse</i>), かちんこちん (<i>kachin-kochin</i>), サンドイッチ状の (<i>like a sandwich</i>)
New food expression	11	Shown in Table 5

Table 5. Candidates of the obtained new food expressions and its example with the originally intended food expression.

Index	Food expression	Examples	The intended food expression
1	分厚い (thick)	さとみんみんさん、分厚いの大好きです。 (<i>Satominmin</i> , I like thick food very much.)	厚い (thick)
2	こってり (fatty)	カレーともやしこってりして美味しい! ^m (<i>Curry</i> and sprout are good tastes because they are fatty)	しこっ (<i>shikox</i>)
3	ジューシー (juicy)	胸肉なのにジューシーで美味しかった☆これから揚げは胸肉で (It is good with meat juice even if it's breast. I'll cook fried chicken with breast.)	からから (<i>kara-kara</i>)
4	歯にくっつく (stick to teeth)	食べたときに歯にくっつく心配ありますが、青のり多目が◎ですね (I wonder it stick to teeth, but a lot of green laver is OK.)	くっつく (stick)
5	しゃくしゃく (<i>syaku-syaku</i>)	しゃくしゃくしゃく…箸休めのはずが、休まりません ^{^^} ; (<i>Syaku-syaku, syaku-syaku, syaku-syaku</i> , I can't stop eating.)	くしゃくしゃ (<i>kusya-kusya</i>)
6	しゃきしゃき (<i>syaki-syaki</i>)	じとじとふきとぼせ! しゃきしゃきびり辛あつというまにたべちゃった (Kick out <i>jito-jito</i> ! I ate the spicy <i>syaki-syaki</i> in a minute.)	じとじと (<i>jito-jito</i>)
7	しゃっきり (<i>syakkiri</i>)	ほんと〜すっごくしゃっきり仕上がってとっても美味しそう (It really should be very yummy with very <i>syakkiri</i>)	くしゃっ (<i>kusyax</i>)
8	ざっぱり (<i>Sappari</i>)	ズッキーニの代わりにしめじを IN! 梅味がざっぱりでおいし〜しめじとっても美味しそう! (Drop <i>Shimeji</i> instead of <i>zucchini</i> ! Apricot flavor is <i>Sappari</i> and yummy, <i>Shimeji</i> must be yummy!)	じとっ (<i>jitox</i>)
9	サクサク (<i>saku-saku</i>)	自己流でぐしゃっとなっていました! がやっとなサクサクに出来ました☆ (Finally I make it with <i>saku-saku</i> , it was <i>gusyax</i> in my own style)	ぐしゃっ (<i>gusyax</i>)
10	ふわふわ (<i>fuwa-fuwa</i>)	ゆちこちこちこさん☆ふわふわ優しい甘さで幸福な気分です o(ε)o (<i>Yukikochikochiko</i> ☆ I'm very happy with <i>fuwa-fuwa</i> sweet)	こちこち (<i>kochi-kochi</i>)
11	ポリッ (<i>porix</i>)	強靱な顎の持ち主ですね! 顎関節症の私はカクカクしながらポリッ (You have a strong jaws! I have jaw arthritis so I eat it like <i>porix</i> and <i>kaku-kaku</i>)	強靱な (<i>tough</i>)

be not for "baked cookie." In this paper, the participants evaluated only sentence without any review information. Thus, the participants should not recognize this sentence was from the review for recipe of "baked cookie." For more sure collecting framework, a picture or some other information of review should be shown in the interaction. That said, such wrong case were few, so this paper assumed the examples that were accepted based on lax criterion as correct examples for food expressions.

4.2.2. Food expressions for which wrong partial-matched examples were obtained

For 19 food expressions which example is shown in the second row in Table 4, all of their examples obtained with partial-match were rejected. The 19 food expressions were generally not used to express for looking and eating. Thus, it was ever difficult to collect examples for such expressions. The current proposed framework shows five candidates to the user. If correct example exists in the database, it would be possible to collect such correct examples for the expressions by increasing the number of presented candidates; but, this modification would increase the efforts of the users. Improvement of how to collect examples for such food expressions (i.e., food expressions that are generally not used to express looking and eating) will be our future work.

4.2.3. Food expressions for which no example was obtained

For 35 food expressions, no examples was obtained with partial-match. The expressions were the either of follows;

Pattern 1 The food expressions are conjugatable.

Pattern 2 Onomatopoeia written in *Hiragana*

Examples of the pattern 1 are “かどぼった (angular),” “口あたりがよい (mild),” and “粗い (rough).” Such expressions were stored in the prepared dictionary with their original form. However, the original form rather be unusually used in casual documents such as recipe reviews. The expressions are usually written as “かどぼった → かどぼっていて,” “口あたりがよい → 口あたりがよかった,” and “粗い (rough) → 粗く.” Such conjugatable expressions can be not obtained with partial-match. More example candidates would be obtained using the partial-match with the fixed part of expressions.

As examples of pattern 2 are “がさがさ (gasa-gasa),” “じとっ (jito),” and “かちんこちん (katin-kotin).” Japanese have three types of characters *Hiragana*, *Katakana* and *Kanji*. *Hiragana* and *Katakana* represent just a pronunciation, and *Kanji* has pronunciation and its own meaning. Native Japanese uses the three characters depending on domain, origin of a word, and affection. Onomatopoeia can be written either *Hiragana* and *Katakana* form. In most case of writing onomatopoeia, *Katakana* form is commonly used more than *Hiragana* form. Since only *Hiragana* forms were stored in the prepared dictionary for these expressions, no example was obtained for these expressions. Preparing both *Hiragana* and *Katakana* forms for onomatopoeia, this problem will be easily solved.

4.2.4. New food expressions

From the candidates of new food expressions shown in Table 5, three candidates were submitted by the participants. In this paper such probable candidates are discussed in detail.

The expression 1 is surely appropriate as a food expression and should be included in the dictionary of food expressions. “分厚い” is a word meaning more thick than “厚い (thick).” However the meanings of both words are almost same, affection or sense from the word are slightly different. For a strict text mining for food reviews, “分厚い” should be assumed as a term in morphological analysis.

The expression 4, which is not a word but phrase, should be a food expression. “くつつく (stick)” was stored but “歯にくつつく (stick to teeth)” was not stored in the dictionary. The prepared dictionary considered that “くつつく (sticking)” was enough as a food expression, however, the participants mentioned that “歯に (to teeth)” should be accompanied. Though teeth are the objects to be stucked in most case of reviewing food, there are the other objects, e.g., tongue and hand. This result suggested that “歯にくつつく” would be better as the food expression to clarify the food texture.

The expression 8 “さっぱり (*sappari*),” was also a new food expression. In Japanese, “さっぱり” has so many meaning depending on document domain; in food domain it means refreshing, light taste, and not greasy. In the dictionary, “こってり (*kotteri*)” and “脂っこい (greasy)” were stored as food expressions, which are the opposite words of “さっぱり.” It seemed that “さっぱり (*sappari*)” was omitted in the collecting food expressions.

In this user test, the number of probable new food expressions were few. However, all of them were surely food expressions and effective to describe food taste and texture. And other candidates were not far from appropriate food expression in our heuristics. The effectiveness to obtain new food expressions will be examined through number of user test.

5. Related work

Food and cooking engineering has been researched from various aspects, e.g., physical and its expressions in language. Especially, food taste and texture, i.e. food expressions are the target of such research.

Sagara *et al* researched the relationships between physical characteristics and expressions in language. They evaluated and digitized food textures with texturometer⁵. The results of their work provides us with the idea that what physical characteristic cases what expression.

Documents in food domain are analyzed in the field of natural language processing and interface. Yamakata *et al* analyzed recipe texts for casual cooking⁶. However, they did not analyze the review for eating food. Tomoto *et al* focused on onomatopoeia, which is often used as food expressions in Japanese, and structured an onomatopoeia thesaurus map⁷. Their map enables us to understand relationships between onomatopoeias for food expressions.

For such document analysis in food domain, Hayakawa *et al* developed a database for food textures⁸. This database was developed based on the subjective evaluation of sensual experts. Their database covers most part of usual food expressions, however, new food expressions often appear and disappear. Thus, the database should be improved; this is the one of the motivation of this paper. The proposed framework enables general users to easily collect food expressions with interaction and defined criteria of food expressions.

6. Conclusions

This paper proposed an interactive framework for collecting correct example sentences for food expressions and discovering new food expressions. Through the user test with the proposed framework, example sentences for 66 food textures were collected. And, three new food expressions were discovered in the user interaction. Through the discussions, the effectiveness of the proposed framework was confirmed.

In our future, example sentences and new food expressions which are obtained with the proposed framework should be evaluated by more users and food professionals. We believe that text mining model for food domain will be improved as using machine learning with the learning data collected through the proposed framework. The effectiveness of the collected data for machine learning will be also experimented in the future.

Acknowledgements

This work was supported in part by JSPS Grant-in-Aid for Scientific Research #16K21482 and Artificial Intelligence Research Promotion Foundation. We are appreciative of the discussions with Prof. Mori at Academic Center for Computing and Media Studies, Kyoto University.

References

1. Shinsuke Mori, Tetsuro Sasada, Yoko Yamakata, and Koichiro Yoshino. A machine learning approach to recipe text processing. In *Cooking with Computers Workshop*, 2012.
2. Ronald Jowitt. The terminology of food texture. *Texture studies*, pages 351–358, 1974.
3. Fumiyo Hayakawa. Development of japanese texture lexicon for food texture studies. *Nippon Shokuhin Kagaku Kogaku Kaishi*, pages 311–322, 2013. (in Japanese).
4. Seiji Yoshikawa. Collection and analysis of texture terms. *Quality control*, pages 66–70, 1968.
5. Yasuyuki Sagara. Kansei measuring methods for mechanical properties and texture of foods. *Nippon Shokuhin Kagaku Kogaku Kaishi*, pages 501–512, 2009. (in Japanese).
6. Yoko Yamakata, Yuichi Sugiyama Shinji Imahori, Shinsuke Mori, and Katsumi Tanaka. Feature extraction and summarization of recipes using flow graph. *Lecture Notes in Computer Science*, pages 241–254, 2013.
7. Yutaro Tomoto, Tsuyoshi Nakamura, Masayoshi Kanoh, and Takanori Komatsu. Visualization of similarity relationships by onomatopoeia thesaurus map. In *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, pages 1–6, 2010.
8. Fumiyo Hayakawa, Yukari Kazami, Katsuyoshi Nishinari, and Kaoru Kohyama. Classification of japanese texture terms. *Texture Studies*, pages 140–159, 2013.