

Explaining Japanese Onomatopoeia in Chinese

Using Translated Paraphrases

Wen Luo, Fumito Masui, Michal Ptaszynski

Kitami Institute of Technology

{luowen,f-masui,michal}@ialab.cs.kitami-it.ac.jp

Abstract As a result of growing popularity of Japanese comic books in China in recent years a large number of such products has been translated into Chinese. However, there has been a problem with translating Japanese comic books, namely, translating Japanese onomatopoeia into Chinese. In most cases Japanese onomatopoeia are either not translated at all, or Chinese characters are used to represent similar pronunciation to Japanese, which is often insufficient for Chinese readers to understand. Thus the Chinese readers often cannot understand the real meaning of Japanese onomatopoeia. This problem is caused by a deep cultural and historical gap between the two languages. Furthermore, meaning and grammar of onomatopoeia has changed reflecting the times, which suggests that updating onomatopoeia dictionaries is an important and challenging task. In this paper, to solve the abovementioned translation problem, we propose a method to translate paraphrases of Japanese onomatopoeia to Chinese for machine translation of onomatopoeia. In particular, we focus on nouns and verbs as key points to describe the meaning of onomatopoeia without literal translation. The success of this study will be play a significant role for the translators and the Chinese people interested in Japanese culture and language.

Key words Japanese onomatopoeia, explanation generation, translated paraphrases, machine translation

1 Introduction

Japanese has a large inventory of sound symbolic or mimetic words, known in linguistics as onomatopoeia. This is one of the distinctive features of the Japanese language. Japanese onomatopoeia consist of onomatopoeia and mimetic expressions. Onomatopoeia refer to the words that mimic actual sounds or voices. Mimetic expressions refer to the words that depict

non-auditory senses, psychological states or bodily feelings. Onomatopoeia contain about 4500 words in Japanese. Because onomatopoeia has vivid expressiveness, it has been widely used in comic books, novels and other literature. It is indispensable device in Japanese everyday language use [1].

The words equivalent to Japanese onomatopoeia are the onomatopoeia in Chinese. However, there have been only few Chinese onomatopoeia when compared with Japanese onomatopoeia, thus it is not possible to translate them directly. Furthermore, onomatopoeia have plural meaning in Chinese. If there is no clear context, Chinese readers often cannot guess the true meaning of onomatopoeia. For example, the Chinese onomatopoeia *wang wang* (汪汪) has two meanings. One is the sound of dog barking, and the other one is a name of a pop-idol. Saying only “wang wang” in a dialogue or writing will not be understood in most cases. Furthermore, the words equivalent to Japanese mimetic expressions existed in Chinese in the past, but they are rarely used today. Shuxiang Lü (吕叔湘) who is famous for writing a Chinese dictionaries, such as *Xiandai Hanyu Cidian*¹ defines them as “vivid adjectives” in his dictionary (*Xiandai Hanyu Babai Ci*) [2]. However, the concept of mimesis has disappeared from Modern Chinese. Now, although onomatopoeia and “vivid adjectives” are used in children’s books and novels, their number is limited. There are only 174 onomatopoeia and vivid adjectives recorded in the dictionary (*Xiandai Hanyu Cidian*) [3]. These facts show that, Chinese onomatopoeia and vivid adjectives comparing with Japanese onomatopoeia, are of limited number and narrow range.

In recent years, in works translated from Japanese

¹ Shuxiang Lü, et al. (2005). *Xiandai Hanyu Cidian* [Dictionary of Chinese] (现代汉语词典). The Commercial Press.

into Chinese, Japanese onomatopoeia are usually not translated at all with the original Japanese onomatopoeia used in the translated works. Another way of dealing with Japanese onomatopoeia is to find a word with similar pronunciation in Chinese instead of actually translating Japanese onomatopoeia. These two translation methods exist in translations. Unfortunately, both cause the meaning of Japanese onomatopoeia to be difficult to understand by Chinese readers.

The purpose of this study is to help Chinese readers understand the actual meaning of Japanese onomatopoeia. To achieve this goal we develop a system for automatic machine translation of onomatopoeia, it can explain Japanese onomatopoeia in Chinese using translated paraphrases.

2 Onomatopoeia

2.1 Japanese onomatopoeia

Modern estimate of the number of Japanese onomatopoeia indicates about 4500 words [4]. This refers to both onomatopoeia and mimetic expressions. Onomatopoeia refer to the words that mimic actual sounds or voices. For example, *shitoshito* (シトシト) conveys the meaning that gentle rain is falling; *dokidoki* (ドキドキ) means that someone's heart is pounding; *wanwan* (ワンワン) means that a dog barks ("bow-wow" in English). Mimetic expressions refer to the words that depict non-auditory senses, psychological states or bodily feelings. For example, *tappuri* (たっぷり) means that there is plenty of something; *uruuru* (うるうる) means that someone's eyes are wet with tears. Because these onomatopoeia are vividly expressive, they have been widely used in comic books, novels and other literature. It is an indispensable device in Japanese everyday language use.

2.2 Chinese onomatopoeia

Chinese onomatopoeia consist of onomatopoeia and vivid interjections. Onomatopoeia refer to the words that

mimic actual sounds. This definition is similar to Japanese onomatopoeia. For example, *wang wang* (汪汪) means that a dog barks (bow-wow). Vivid interjections refer to the words that are used to express emotions or sentiments. For example, *yi* (咦) is an expression of surprise. The words equivalent to Japanese mimetic expressions existed in the past. However, they are rarely used today. Furthermore, the concept of mimetic expressions has disappeared from Modern Chinese. There are only 174 containing onomatopoeia and vivid adjectives recorded in the dictionary (*Xiandai Hanyu Cidian*). Now, although onomatopoeia and the vivid adjectives are used in children's books and novels, they are usually neglected in everyday language use in Chinese.

3 How to understand Chinese onomatopoeia

The problem we tackle in this research is that Chinese readers cannot understand the true meaning of Japanese onomatopoeia, even when they read translated works. We assume that the issue arises from the fact that improper translation method is applied in translating Japanese comic books to Chinese.

One way of dealing with Japanese onomatopoeia in Chinese translations is leaving the original Japanese onomatopoeia untranslated. Translated works simply keep the original characters in the translation, as shown in Figure 1. This makes Chinese readers unable to understand the meaning of Japanese onomatopoeia. A second way of dealing with Japanese onomatopoeia is that the translator uses Chinese words with the pronunciation similar to the Japanese words instead of actually translating the Japanese onomatopoeia. An example is represented in Figure 2, where the word *gaa* (ガア) in the original is translated into *gua ya* (呱呀) in Chinese. Unfortunately these characters are not defined in Chinese dictionary and thus it is difficult to understand for Chinese readers.



Figure 1 Example of leaving Japanese onomatopoeia untranslated

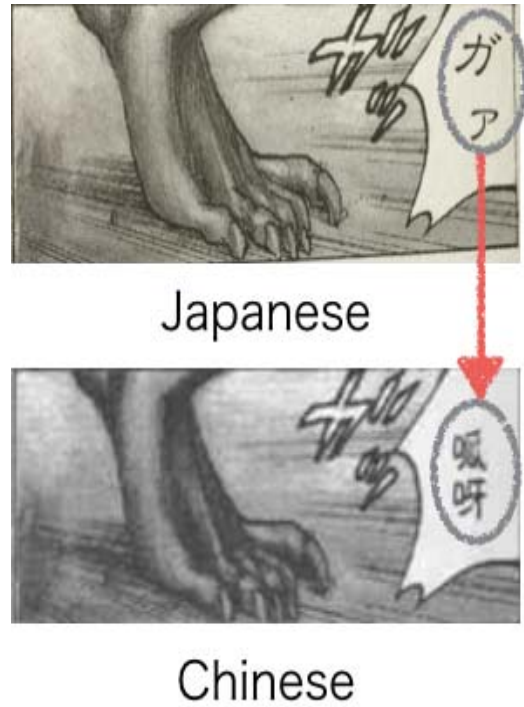


Figure 2 Example of translating onomatopoeia using unrelated characters

4 Proposed method

In order to solve the above problem we develop BeiSen-word a system for machine translation of onomatopoeia, which translates paraphrases of Japanese onomatopoeia to Chinese. Summary of the system is shown in Figure 3. BeiSen-word obtains query x_i from the user containing onomatopoeia, then uses

metaphorical paraphrases of the onomatopoeia generated with the method developed by Masui et al. (2012) [5] to search for hints that suggests the meaning of the onomatopoeia in the Web, and generates the implied sentences in Japanese. Then BeiSen-word translates the implied sentences into Chinese and returns the output to the user. The configuration diagram of the system is shown in Figure 4.

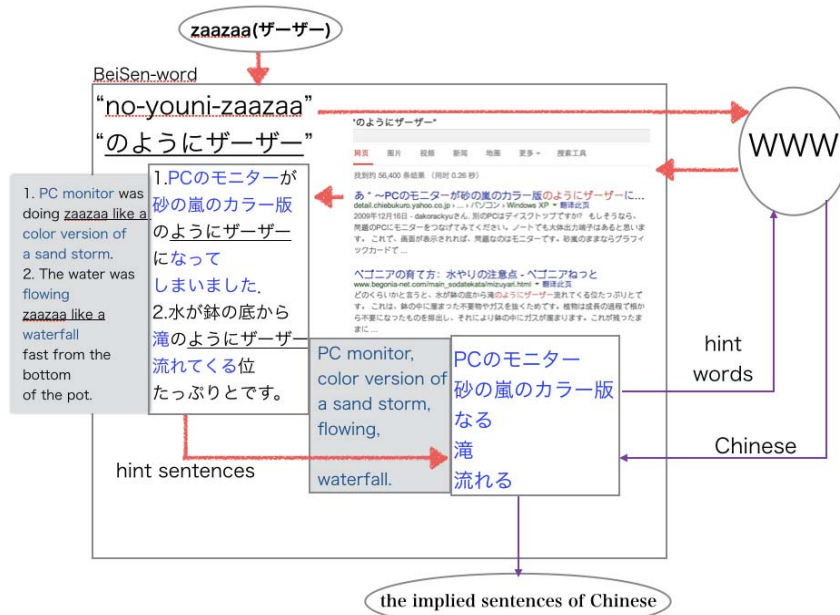


Figure 3 An example showing the processing follow of the method

Below we describe BeiSen-word's translation process divided into 7 steps.

(STEP 1) Obtaining the query. In this step the query is obtained from the user. It is identified as x_i , and it is divided into onomatopoeia, mimetic expressions and unknown after text segmentation.

(STEP 2) Generation of metaphor phrases. Preparing metaphor phrases consisting of *no-youni* (のように) and *no-youna* (のような). When the query x_i is identified as onomatopoeia, it is combined in the form *no-youni- x_i* (のように x_i); When the query x_i is identified as mimetic expressions, it is combined in the form *no-youna- x_i* (のような x_i); When the query x_i has been identified as unknown, it is expressed as *no-youni- x_i* (のように x_i) and *no-youna- x_i* (のような x_i).

(STEP 3) Search. The *no-youni- x_i* (のように x_i) and

no-youna- x_i (のような x_i) phrases are used as query to a search engine.

(STEP 4) Hint snippet extraction. In this step snippets containing the Japanese sentences with the queried onomatopoeia are extracted from search engine results.

(STEP 5) Hint word extraction. Nouns, noun phrases, verbs and verb phrases that describe (surround) the queried onomatopoeia are extracted from the snippets as hint words for further processing.

(STEP 6) Translation of hint words. Extracted words and phrases are translated using widely available translation online service.

(STEP 7) Generation of translation. The translated words and phrases are used to produce descriptions of Japanese onomatopoeia in Chinese.

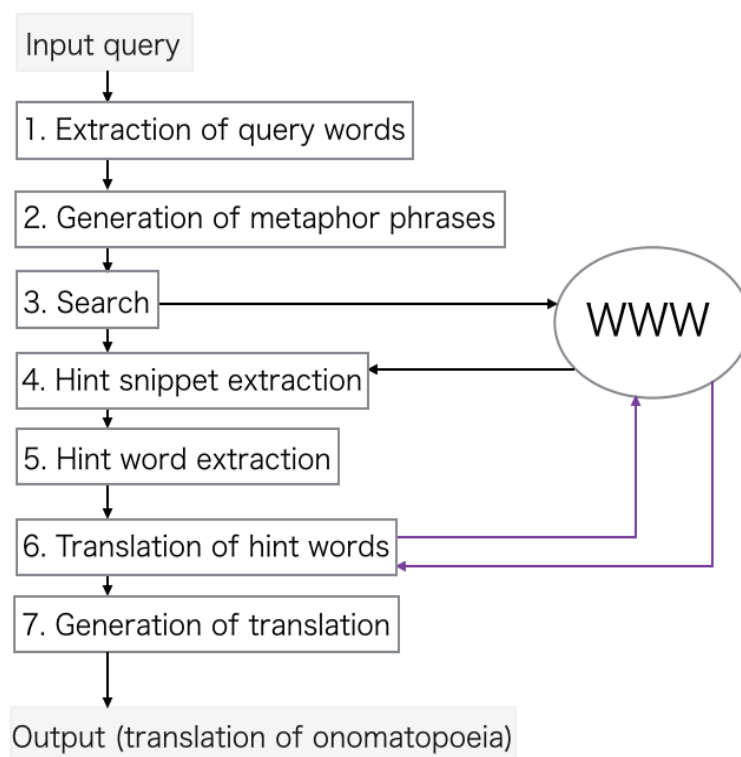


Figure 4 The flow chart of BeiSen-word's translation process divided into 7 steps

5 Preliminary investigation

In this study we generate translated paraphrases of Japanese onomatopoeia using simplified sentence templates constituted by nouns and verbs. In order to determine the sentence elements in Chinese, we investigate two aspects of how Chinese people use onomatopoeia in daily life. Both aspects will be discussed below.

5.1 Investigation of morphological patterns of onomatopoeia explanations in Chinese dictionary

We investigated the contents of Chinese dictionary in order to find out how onomatopoeia are explained in Chinese. We extracted 40 groups of onomatopoeias with their explanations from the *Xinhua Zidian* (新华字典) dictionary.

From those explanations we were able to extract

seven kinds of part-of-speech (POS) morphological patterns. These patterns are shown in Table 1. The table lists POS patterns of Chinese onomatopoeia, their frequency and examples of their use. Furthermore, within

the 40 samples there were 9 groups of onomatopoeia which had no explanation. We define them as exceptions.

Table 1 Morphological patterns of onomatopoeia explanations in Chinese

POS pattern	Frequency	Examples	
		Onomatopoeia	Meaning in Chinese
n + de shengyin (noun + 的声音, the sound of + noun)	12	huhu (呼呼)	fengsheng (风声, sound of the wind)
n + v + de shengyin (noun + verb + 的声音)	13	peng (澎)	bolang chongjisheng (波浪冲激声, sound of the wave)
adj + de shengyin (adjective + 的声音)	1	wuwu (呜呜)	dichen de shengxiang (低沉的声响, grave sound)
v + de shengyin (verb + 的声音)	1	pipa (哧啪)	paida de shengyin (拍打的声音, the sound when tapping)
v + n + de shengyin (verb + noun + 的声音)	2	jiangjiang (将将)	jinshi zhuangji de shengyin (金石撞击的声音, the sound when the stones hit)
adv + v + de shengyin (adverb + verb + 的声音)	1	longlong (隆隆)	julie zhendong de shengyin (剧烈震动的声音, the sound when severe shock)
adj + v + de shengyin (adjective + verb + 的声音)	1	jiujiu (啾啾)	qili de jiaosheng (凄厉的叫声, the sound of shrill cry)
Exceptions	9	dongdong (咚咚)	-
Sum	40	-	-

As the result of investigation, the pattern *noun + de shengyin* (noun + 的声音, “the sound of + noun” in English) POS pattern accounted for 30.0% of all; *noun + verb + de shengyin* (noun + verb + 的声音) accounted for 32.5%; only these both together accounted for 62.5%. This shows that, this two kinds of POS pattern have a dominant position in interpretation of Chinese onomatopoeia.

5.2 Psychological experiment with onomatopoeia explanations in video samples

We conducted a psychological experiment to investigate how Chinese people understand Japanese

onomatopoeia.

The methodology of the experiment was as follows. We let four Chinese students from different residence watch a video, and describe the onomatopoeia explanation scene in the video in Chinese. This video constituted of 15 onomatopoeia scenes, each onomatopoeia being a sound from daily life. For example, the sound of an oven when finished making a toast in Japanese is *chin* (ちん). In the video a toaster making a toast and the accompanying sound *chin* (ちん) are represented together in the same scene, as shown in Figure 5.



Figure 5 One picture from the video explaining the sound of oven making a toast

The results of the experiment are shown in Table 2. For example, the four Chinese people who saw the scene explaining onomatopoeia *chin* (ちん) responded the following ways:

1. *ding* (叮, “the sound of *ding*”)
2. *ding* (叮, “the sound of *ding*”)
3. *mianbao kao haole* (面包烤好了, “Bread baked.”)
4. *ding de yisheng and mianbao kao haole* (叮的一声, 面包烤好了, “Bread baked when the sound

of *ding* from the oven.”)

As a response to the scene explaining the onomatopoeia *peta* (ぺた), the participants responded the following ways:

1. *pa* (啪, “the sound of *pa*”)
2. *da* (嗒, “the sound of *da*”)
3. *nianzhu* (粘住, “stick”)
- tie* (贴, “post”)

Table 2 Details of the experiment using onomatopoeia videos with details of the respondents

Video title	<i>Sotugyou seisaku onomatope no aru seikatsu</i> (卒業制作オノマトペのある生活, graduate work: Onomatopoeia in life)			
Date of experiment	2014/4/22			
Participant	1	2	3	4
Place of birth	广东省 (Guandong Province)	云南省 (Yunnan Province)	四川省 (Sichuan Province)	河南省 (Henan Province)
Language proficiency	Chinese and Japanese	Chinese and Japanese	Chinese and English	Chinese

As the result of the experiment each of the four interviewed respondents produced explanations in Chinese of 15 onomatopoeia videos. Three people translated all the onomatopoeia into Chinese in the video, and one person failed to translate two onomatopoeias

from the videos. In their answers we can see one of the methods of dealing with Japanese onomatopoeia mentioned in the Introduction, namely, to find a word with similar pronunciation in Chinese instead actually translating the Japanese onomatopoeia.

In Table 3, the explanation without any marks can be easily understood by other Chinese people. On the other hand, the explanations in marked with † and ‡ cannot be understood by other people. These answers in Chinese were compiled as shown in Table 3. Then they were showed other Chinese people, to test whether these

explanations can be understood by other Chinese people. There were 17 groups which could not be understood by other people. Therefore these 17 groups were not included in the analysis statistics. We analyzed the remaining 43 groups and checked POS patterns they conatined. The results are shown in Table 4.

Table 3 Results of the experiment using onomatopoeia videos

Onomatopoeia from video	Participant 1	Participant 2	Participant 3	Participant 4
kachi (カチ)	†哒哒哒 (the sound of dadada)	†喀喀喀 (the sound of kakaka)	滴答地转 (tick of the clock)	滴滴地转 (tick of the clock)
chin (ちん)	†叮 (the sound of ding)	†叮 (the sound of ding)	面包烤好了 (Bread baked.)	叮的一声，面包烤好了 (Bread baked when the sound of ding from the oven.)
karakara (からから)	空杯 (empty cup)	啥都没有 (nothing)	空杯 (empty cup)	空杯 (empty cup)
chobi (ちよび)	四分之一 $\frac{1}{4}$	‡一丢丝 (a little bit)	四分之一 $\frac{1}{4}$	一点点 (a little bit)
sokosoko (そこそこ)	半杯 $\frac{1}{2}$ cup)	半杯($\frac{1}{2}$ cup)	半杯 $\frac{1}{2}$ cup)	装满一半 (filled half)
choudo (ちょうど)	三分之二 $\frac{2}{3}$	半杯 $\frac{1}{2}$ cup)	一杯 (one cup)	快要满了 (It is almost full.)
naminami (なみなみ)	满杯 (cupful)	一大杯 (a large cup)	满杯 (cupful)	满了 (full)
parapara (ぱらぱら)	撒盐 (sprinkle salt)	†撒撒撒 (the sound of sasasa)	沙沙地撒 (the sound of shasha when sprinkle salt)	撒盐 (sprinkle salt)
jiwa (ジワ)	†吱 (the sound of zhi)	化掉 (burn off)	融化了 (melted)	融掉了 (melted)
saku (さく)	-	‡咔吡 (the sound of kaci)	啃面包 (eat bread)	他把面包吃掉了 (He ate the bread.)
kirakira (きらきら)	飘落 (falling)	飘 (gone with the wind)	飘落 (falling)	雪花般飘落 (snowflakes falling)
peta (ぺた)	†啪 (the sound of pa)	†嗒 (the sound of pa)	黏住 (stick)	贴 (post)
tantan (たんたん)	†嗒嗒嗒嗒 (the sound of dadadada)	†噼里啪啦 (the sound of pilipala)	打字 (typing)	打字 (typing)
biri (びり)	-	‡吡啦 (the sound of cila)	打开文具盒 (open cases)	打开文具袋 (open stationery bag)
guxtu (ぐつ)	†嘣 (the sound of beng)	†杠 (the sound of gang)	当地盖章 (the sound of dang when cover seal)	盖印章 (cover seal)

Table 4 POS patterns that appeared in comprehensible explanations of onomatopoeia videos

POS pattern	Frequency
n	20
v	4
v + n	6
v + past form (了)	4
adj + n	1
onomatopoeia + v	3
adv	1
sentence	3
Sum	43

6 Results

In this section we explain and discuss the results of the two experiments.

6.1 Result of the experiment investigating the meaning and POS patterns in onomatopoeia included in Chinese dictionary

As we investigated the explanations of onomatopoeia in Chinese dictionary we found out that there were nine onomatopoeias that had no specific meaning. They are included in the dictionary as onomatopoeia, however, the sound they represent can be construed differently depending on a situation. Onomatopoeia like these, although they are included as onomatopoeia in the dictionary, because they do not represent any particular sound we consider them in our research as exceptions.

In Table 1, the POS pattern *noun + deshengyin* (noun + 的声音) and *noun + verb + deshengyin* (noun + verbs + 的声音) contain a total of 25, which accounts for 80.6% of onomatopoeia after removing the exceptions. From this we say that focusing only the noun and verb POS patterns should be sufficient in our method.

Furthermore, speech pattern *noun + deshengyin* (noun + 的声音) is a form of the *noun + verb + deshengyin* (noun + verbs + 的声音) pattern when the verb is omitted. It is common in Chinese to avoid long explanations, when explaining everyday commonsense meanings. Explanations in dictionaries are also often abbreviated. Therefore we assume that the meaning of onomatopoeias can also be briefly explained in Chinese

with only those simple patterns.

For many words, if the explanations are too long some parts can be omitted, while still resulting in correct interpretation in Chinese. For example, *huhu* (呼呼) is interpreted as *fengsheng* (风声, “sound of the wind”) in Chinese. The word *fengsheng* (风声) when directly translated into Japanese becomes *kaze no oto* (風の音, “sound of the wind”). A Japanese verb *fuku* (吹く, “to blow”) can be incorporated in this interpretation, which results in a new interpretation, also comprehensible in Chinese, namely, *feng chui de shengyin* (风吹的声音, “the sound of blowing wind”). Thus in cases where the verb is not omitted in the dictionary of Japanese onomatopoeia, the verbs can also be retained when translating the onomatopoeia from Japanese to Chinese. There is also no need to change back the *noun + deshengyin* (noun + 的声音) pattern into *noun + verbs + deshengyin* (noun + verbs + 的声音).

6.2 Results of psychological experiment with onomatopoeia explanations in video samples

In Table 4 we can see that onomatopoeia are most often interpreted in Chinese using nouns, which accounts for 46.5%. Thus we can say that noun based patterns are very important in generation of the explanations.

In the experiment, the fact that the participants were originally from different places within China, caused dialectal differences in some of the translated interpretations. For example, one participant translated onomatopoeia such as *chobi* (ちよび), *saku* (さく) and *biri* (びり) into dialect words of Qujing city in Yunnan province. These translations can be understood only people who are familiar with this dialect. This shows that the generated translation sentences should be generated in Mandarin Chinese for them to be understandable by all Chinese readers.

On the other hand, two of the subjects know Japanese. They translated Japanese onomatopoeia into Chinese with similar pronunciation after watching the video. When the translations were shown to other 10 Chinese people, none of them was able to understand the meaning of the onomatopoeia. This is because the translations are not standard Chinese words (newly coined words) and without a specific context it is not possible to understand the meaning of those words. Thus these answers are marked in Table 3 with † as not usable for our purposes. In summary, comprehensibility of

translating Japanese onomatopoeia into words with similar sound is very low.

Looking at Table 3, we can see that subject 1 was unable to translate the onomatopoeia *saku* (さく, “crunchy, crispy”). The reason for this can be found in cultural difference between Japan and China. In Japan there is a habit of eating baked bread. However, in China there is no habit to eat baked bread, thus subject 1 was not able herself to understand that eating freshly baked bread results in occurring a crispy sound. This shows that the process of translation of onomatopoeia could be prone to revealing such cultural differences. Therefore it would be necessary to explain the cultural differences in addition to the translation.

In addition, subject 1 was also not able to translate the word *biri* (びり, “sound of tearing up a piece of paper”). As the reason for this it could be considered that since the number of onomatopoeia is small in the Chinese it was difficult for the participant to imagine a Chinese equivalent of the Japanese onomatopoeia. This shows that it is necessary to provide support for Chinese readers in understanding Japanese onomatopoeia.

7 Conclusion

The goal of this study is to provide Chinese readers with means to understand the actual meaning of Japanese onomatopoeia. To achieve this goal we conducted a number of studies to investigate ways both Japanese and Chinese express and understand onomatopoeia. This knowledge will be used to develop a system for automatic machine translation of onomatopoeia, which will explain Japanese onomatopoeia in Chinese using translated paraphrases. In order to determine the elements of sentences used in explanation of onomatopoeia in Chinese, we investigated two aspects of

how Chinese people use onomatopoeia in daily life. We found out that Japanese onomatopoeia translated into Chinese with the use of POS pattern consisting of nouns, verbs and nouns + verbs is sufficient to facilitate the meaning of Japanese onomatopoeia in Chinese readers.

In the near future we plan to conduct a deeper analysis of this finding. We will translate Japanese onomatopoeia using the above POS patterns (nouns, verbs and nouns + verbs) to see whether such simplified explanations can actually be understood by Chinese readers.

References

- [1] Yuzu Uchida, Kenji Araki, and Jun Yoneyama, (2012). *Burogu kiji kara no onomatope youreibun no jidou chuushutu shuhou* [A Method of Onomatopoeia Sentences Extraction from Blog Entries] (in Japanese), *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, 24.3 (2012): 811-820.
- [2] Lee Kyeong-A, (2007). *Xiandai hanyu nishengci yanjiu* [Onomatopoeia in Modern Chinese] (in Chinese), Shanghai, Xuelin chubanshe.
- [3] Renying Dayan, (2006). *Dangdai hanrinishengci duibi yanjiu* [A Comparative Study of Onomatopoeia in Chinese and Japanese] (in Chinese), Fudan University.
- [4] Masahiro Ono, (2007), *Giongo, gitaigo 4500 nihongo onomatope jiten* [Dictionary of Japanese onomatopoeia] (in Japanese). Shogakukan.
- [5] Fumito Masui, et al. (2010). *WWW hatsuyou ni yoru go no hiyuteki sobyou shuhou* [A method for generating metaphorical descriptions using the World Wide Web] (in Japanese), *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics* 22.6 (2010): 707-719.
- [6] Wanhui Feng, (2011). *Xinhua Zidian* [Dictionary of Chinese] (in Chinese). The Commercial Press.