

Onomatopen: painting using onomatopoeia

Keisuke Kambara and Koji Tsukada

Ochanomizu University, Ochadai Academic Production,
2-1-1 Otsuka, Bunkyo-ku, Tokyo, Japan
kambara@sappari.org, tsuka@ubiquitous.com
+81-3-5978-2571
<http://sappari.org/>

Abstract. We propose an interactive technique using onomatopoeia. Onomatopoeia are imitative words such as “Zig-zag” and “Tick-tock”. Some Asian languages, especially Japanese and Korean, have many onomatopoeia words, which are frequently used in ordinary conversation, as well as in the written language. Almost all onomatopoeic words represent the texture of materials, the state of things and emotions. We consider that onomatopoeia allows users to effectively communicate sensory information to a computer. We developed a prototype painting system called Onomatopen, which enables a user to switch brushes and apply effects using onomatopoeia. For example, if the user draws a line while saying “Zig-zag Zig-zag...”, a jagged line is drawn. As a result of our user test, we found that users can easily understand the usage and enjoy drawing with this application more than with conventional painting software.

Key words: Onomatopoeia, Multimodal input, Pen-based interface, Voice-based interface

1 Introduction

Onomatopoeia is a general expression used in ordinary spoken and written language. Some kinds of onomatopoeic words imitate sounds, such as the sound of a clock: “Tick-tock”. Others mimic states or emotions, such as “Zig-zag”.

The use of onomatopoeia varies with language, and some Asian languages, especially Japanese and Korean, have many onomatopoeia words.

Onomatopoeic words represent states, movements, feelings and emotions, and allow their expression in a fun, lively manner. Children typically call cars “Zoom” or call dogs “Bow-wow”, as onomatopoeic words are easy to pronounce and remember even for children.

Although onomatopoeic words are often used in conversation and writing, they are not used in human-computer interactions. Among the many research projects on interactive techniques using voice input, few research projects have focused on onomatopoeia. Herein, We propose a friendly, sensory-based interactive technique based on the use of onomatopoeia.

2 Onomatopen

We applied the interactive technique using onomatopoeia to painting software, and developed a prototype called Onomatopen. Onomatopen enables users to draw textured lines using onomatopoeia, and the texture is associated with the onomatopoeic word spoken. In case of drawing a dotted line, the user draws a line while saying “Ten Ten...¹”, and the texture of the line becomes dotted (Fig.1). In a similar way, using various onomatopoeia, the user can control certain functions of painting software such as switching brushes and applying effects to images.

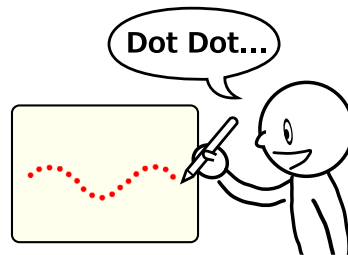


Fig. 1. Concept Image of Onomatopen

2.1 Features

The novel features of Onomatopen compared with existing painting software are as follows.

- Suitable mapping between operation methods and functions
- Multimodal interaction
- Entertainment applications

Suitable mapping between operation methods and functions Onomatopoeia is suitable for representing states of objects in a more sensory-based manner. In Onomatopen, users can imagine the texture of a line or executed function more easily than by changing brushes using buttons.

It is important to associate operation methods with functions intuitively in painting software, because drawing pictures requires sensitivity and feeling.

Multimodal interaction One of the important features of Onomatopen is the ability to draw lines using the voice. This is a kind of multimodal interface. We can help users input more details, such as texture, while drawing. Furthermore, this interaction omits explicit switching operations such as selecting a brush.

¹ “Ten” means a dot in Japanese.

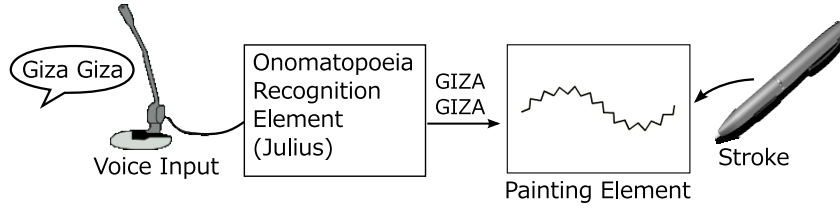


Fig. 2. System architecture

Entertainment applications The above features of Onomatopen, such as appealing to sensitivity and visual/auditory feedback, are suited to the entertainment of the user. Users enjoy synchronizing the pen input and visual/auditory feedback. Drawing for play or as a hobby needs to be fun not only efficient.

2.2 Functions of Onomatopen

We have currently implemented 10 features using onomatopoeia (Table 1, Fig.3). There are two main uses of onomatopoeia; switching brushes and editing images.

Table 1. Functions corresponding to onomatopoeia

Onomatopoeia	Function (Edit, Effect)
Keshi Keshi (means erasing)	Eraser
Choki Choki (means cutting)	Cut
Peta Peta (means pasting)	Paste
Copy Copy	Copy
Biyoon (means deformation)	Deformation

Switching brushes In the case of general painting software, brushes are switched using menus or keyboard shortcuts. On the other hand, users can change brushes simply by saying the appropriate onomatopoeia for the desired texture in Onomatopen (Fig.1). For example, by drawing while repeatedly saying “Zig-zag Zig-zag...”, the texture of the brush automatically switches to Zig-zag. Figure 3 shows the brushes available and the corresponding onomatopoeia.

Editing images We have also developed editing functions such as erasing and cut/paste using onomatopoeia.

Users can erase drawn pictures by stroking the picture while saying “Keshi Keshi...”² (Fig.4).

² “Keshi” means erase in Japanese.



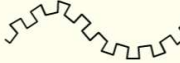


Brush	Onomatopoeia In Japanese	In English
	Giza Giza	Zig-Zag
	Ten Ten	Dot Dot
	Kaku Kaku	Square Square
	Moko Moko	Lumpy Lumpy
	Kira Kira	Twinkle Twinkle

Fig. 3. Brushes corresponding to each onomatopoeia

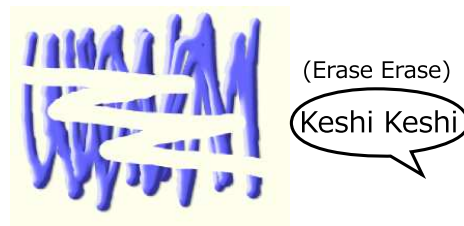


Fig. 4. Erase function

We have also developed copy/cut/paste functions using onomatopoeia. Pictures are cut by stroking while saying “Choki Choki...”³ and copied by stroking with saying “Copy Copy”. Cut or copied pictures are pasted by clicking while saying “Peta”⁴. By stroking with saying “Peta Peta...”, the picture is pasted repeatedly and acts like a kind of brush (Fig.5).

Deformation Effect We have developed a function to apply a deformation effect using onomatopoeia. By dragging a picture while saying “Biyoon”⁵, the dragged area is stretched like rubber (Fig.6).

³ “Choki” means snip in Japanese.

⁴ “Peta” means pasting in Japanese.

⁵ “Biyoon” represents the sound of stretching a rubber-like material in Japanese.

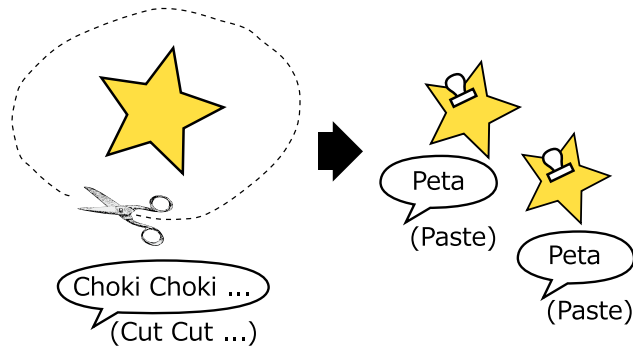


Fig. 5. Cut & Paste

3 Implementation

The Onomatopen system consists of two parts; the voice recognition element and the painting element (Fig.2). The voice recognition element extracts onomatopoeia from user's voice in real-time. The painting element receives the results from the recognition element and switches to the function corresponding to the onomatopoeia.

3.1 Voice Recognition Element

The voice recognition element is based on Julius[1], which is an open-source voice recognition engine. Julius allows the user to customize the grammar for recognition. Onomatopen uses the grammar that recognizes repeated onomatopoeia. Figure 7 show the state transition diagram for the recognition of the repeating onomatopoeia "Zig-zag Zig-zag...".

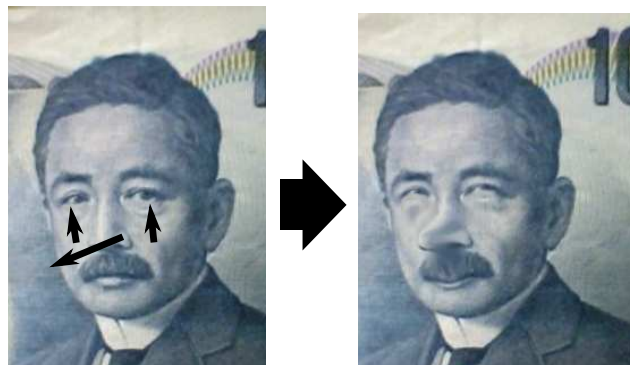


Fig. 6. Deformation Effect

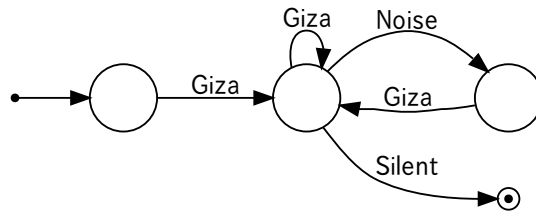


Fig. 7. State transition diagram recognizing “Giza Giza...”

Julius runs as a socket server, which sends the result from the recognition element to the painting element immediately after recognition, the painting element requires the recognition results while the pen is in motion and the onomatopoeic phrase is being repeated.

3.2 Painting Element

The painting element was implemented in Adobe AIR. This element receives the onomatopoeic words from the voice recognition server and switches to the brush corresponding to the onomatopoeia.

It is a burden for the user to say the onomatopoeic phrase repeatedly throughout stroking. To reduce the burden, even after the user stops repeating the phrase, the brush continues as selected until the end of the stroke. The painting element corrects voice recognition errors as described below.

3.3 Voice Recognition Error Correction

The voice recognition system cannot avoid recognition error. If the brush reflected the recognition result without correction, it would frequently switch to an unintended format.

To correct recognition error, we took advantage of the following characteristic of onomatopoeia. In Japanese, almost all onomatopoeia consists of the repetition of the same word. Conversely, it is unusual for other words to be added in repetitions; for example, “Zig-zag Dot Zig-zag...”. When such an unusual pattern is recognized, the onomatopoeia pattern is corrected as a repetition of “Zig-zag”.

Using this method, the brush changes format on the basis of the most frequently appearing word. When the brush changes format during a stroke, the drawn line is rendered again from the start point.

According to this correction method, if an unintended line is drawn, the user can fix it by simply repeating the correct word several times (Fig.8).

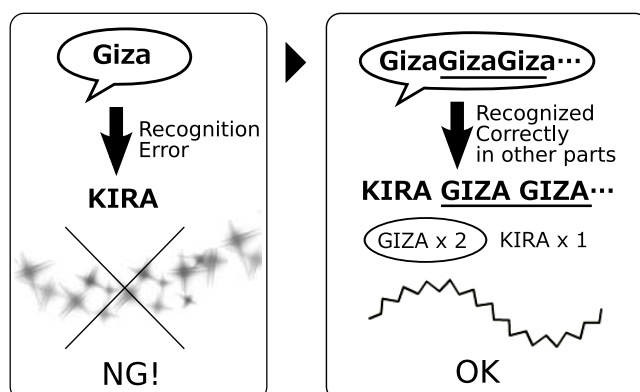


Fig. 8. Error correction by repeating onomatopoeia

3.4 Gap between the beginning of a stroke and the beginning of speech

Oviatt[10] showed that the order and the gap between the beginning of a stroke and the beginning of speech differs according to the individual. If the user occasionally begins repeating the onomatopoeia after the beginning of the stroke, the type of brush can not be defined from the beginning. In such case, a light gray line is drawn at the beginning, and this line is drawn again after the brush format has been defined by voice recognition. The user can draw lines regardless of the gap between the beginning of a stroke and the beginning of speech.

4 Evaluation

We evaluated the prototype by user testing and observation.

4.1 Experiment

After brief instruction regarding the features, each subject freely drew pictures using the Onomatopen for about 10 minutes (Fig.9). Last, we surveyed the subjects by a questionnaire containing questions scored from 1 to 5 and by a free form.

4.2 Subjects

9 subjects (22-32 years old Japanese, 7 females and 2 males) participated in this experiment. The study group contained an expert in painting who uses painting software almost everyday and 3 people who had had little experience with painting software.

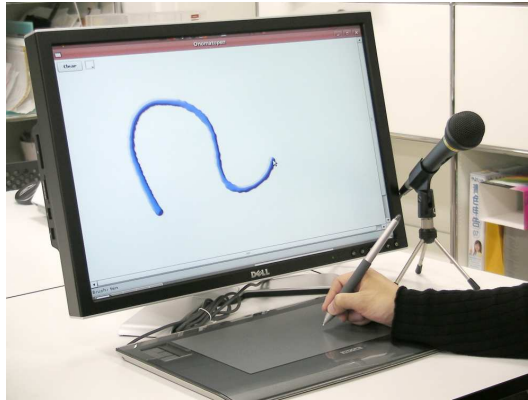


Fig. 9. Evaluation Setting

4.3 Results and Discussion

Tables 2-5 show the results from the questionnaire.

Table 2. Did you feel the mapping between functions and operations natural?

Natural	6	3	0	0	0	Unnatural
---------	---	---	---	---	---	-----------

Table 3. Could you soon understand the usage?

Yes, immediately	8	1	0	0	0	No, not at all
------------------	---	---	---	---	---	----------------

Correspondence between functions and operations As shown in Table 2, all subjects answered that the mapping between functions and operations was “Very Natural” or “Natural”. This result supports the idea that switching the brushes and functions using onomatopoeia is an intuitive interaction.

Eight of 9 subjects could remember the usage immediately (Table 3); therefore, it is reasonable to suppose that the interaction technique is a easy to remember regardless of one’s experience with painting software.

One of the reasons why subjects could remember all the onomatopoeia easily is that the number of onomatopoeic phrases recognized in the current prototype

Table 4. Could you draw as you like without failure?

Yes, perfectly	1	2	2	4	0	No, not at all
----------------	---	---	---	---	---	----------------

Table 5. Could you enjoy drawing?

Enjoyed	7	0	1	1	0	Bored
---------	---	---	---	---	---	-------

is limited in 10. If the number of onomatopoeic phrases is increased, it will be hard to remember those onomatopoeic words. Although this prototype implements brushes and functions that are easy to express in onomatopoeia, some of brushes and functions have no/many corresponding onomatopoeia.

Textures or functions imagined from onomatopoeic phrases differ according to the individual. In case of the experiment, a subject commented that “The lumpy texture is different from what I imagined”. Thus it became clear that brushes implemented in this prototype did not always correspond to what the subjects imagined.

For this reason, the following points require further survey; the kind of brush associated with certain onomatopoeia, the number of onomatopoeia that a user can remember, and individual differences in the understanding of specific from onomatopoeia,

Although it is unrealistic to control all functions in existing painting software using onomatopoeia for the above-mentioned reasons, we suppose it is feasible to assign some functions to specific onomatopoeia as shortcuts. In addition to the normal controls using GUI, by assigning onomatopoeia to particular functions which are frequently used or are easy to express in onomatopoeia, the user will not have to remember all onomatopoeia, and be gradually able to learn the necessary onomatopoeia. If the user can add/change the assignment, individual differences will not be a problem.

Multimodal Interaction The input method of Onomatopen, which uses a pen and voice simultaneously, is not popular in current human-computer interactions. Although all subjects could understand the basic usage immediately, 4 of 9 subjects answered they could not draw as much as they liked (Table 4), and the prototype received low scores. One subject commented that “Pasting is a little difficult to use. Occasionally voice recognition failed, and it did not work correctly.”. Since the onomatopoeia corresponding to the stamp tool and the deformation effect are not pronounced repeatedly, error correction based on onomatopoeia repetition did not work, and failures often occurred when there was a gap between the beginning of the stroke and speech. Another reason for the low evaluation is the lack of an Undo/Redo function. This prototype did not have an Undo/Redo function and the user could not undo a failed stroke. The

problem with the user not being able to draw as much as they liked requires further work and we hope it will be resolved by the implementation of Undo/Redo and the reduction of voice recognition errors.

Error correction by onomatopoeia repetition worked. Subjects who understood the function could draw lines as intended by repeating the onomatopoeic word when the line was incorrect at the beginning of the stroke. Although subjects did not notice the function in the beginning of the experiment, they could quickly understand its use after it was explained and they were able to employ the function.

We plan to carry out quantitative evaluation of the error correction function and detailed analysis of the gap between the stroke and speech in order to improve the multimodal interaction.

Entertainment Seven of 9 subjects felt that drawing using Onomatopen was fun (Table 5), and 5 subjects commented that it was very enjoyable. It is likely that Onomatopen entertains the user more than existing painting software, although there were subjects who could not draw as much as they liked. On the other hand, some subjects commented that “It’s also fun to fail” and “It will be exciting if I draw with many people”. If Onomatopen is used for entertainment, the failures will be simply a part of playing.

5 Possible Applications

Further possible applications of Onomatopen include the following.

5.1 Maps

One merit of Onomatopen is that the user can draw decorated lines quickly and easily. When a user draws a map for guidance, it is better that the lines representing rails, streets and rivers are shown differently. For example, rails and trains are drawn by saying “Clickety-clack”, streets and cars are drawn by saying “Zoom”. In this way, we believe the user can draw maps more efficiently.

5.2 Animation and Games

It may be possible to move still pictures using onomatopoeia representing movement. An airplane can be made to fly by saying “Zoom” or a train made to run by saying “Clickety-clack”.

In the similar way, video games can be controlled by onomatopoeia. For example users can make characters jump by saying “Hop!” or fire a gun by saying “Bang!”.

6 Related Works

6.1 Multimodal interface using voice input

VoicePen[4] proposed a technique to control the thickness, color, transparency and definition of lines using voice characteristics such as loudness and pitch. Although it is similar to Onomatopen in that the textures of lines is controlled by the voice, our research focused on onomatopoeia as information provided by the voice.

Igarashi et al.[5] and Harada et al.[3] proposed interactive techniques using non-verbal information, and reported that non-verbal input enables users to continuously control the strength of effects with resistivity against noises. It will be possible that Onomatopen can adopt such a technique for controlling the thickness and transparency of lines by the volume or pitch of the voice, in addition to the control by onomatopoeia alone.

A multimodal interface that combines pen input and voice commands has been discussed for a long time[2][6]. Onomatopen is different in that it is focused on onomatopoeia for voice input.

Kurihara et al.[8] proposed a system that supports writing on a electronic blackboard during lectures by predictions, and allowing the completion of hand-written notes from the talk. They reported that it was designed to aid writing more quickly, easily and clearly to record information provided in lectures. Onomatopen will also be able to support such situations.

6.2 Sound Symbolism of Onomatopoeia

Onomatopoeia is mainly studied in the fields of literature, linguistics and psychology. A representative study is Tamoris' investigation of the phonological characteristics of onomatopoeia [11]. Murakami[9] analyzed the sound symbolism, which is the relation between the phoneme and the meaning. He extracted phonemic components from Japanese onomatopoeic words by multivariate analysis, and clarified the relation between the phoneme and the meaning in terms of emotions, impressions and materials.

In the field of computer science, Komatsu demonstrated technology to digitize the impressions from onomatopoeia based on the acoustic meaning of the phoneme [7]. Each onomatopoeia is expressed by 8 attribute values such as hardness, smoothness and humidity. We may apply this technology to automatically associate an onomatopoeic word with a brush. Another approach to generate new brushes could be also possible through the analysis of sounds and directly extracting the symbol for the sound without recognizing the spoken word. While the former approach uses only symbolized and simplified information from the sound, the latter approach has an advantage in using much more information from the sound. With the latter approach it may be possible to control the expression based on the way of pronunciation. However, loudness and pitch differ according to the individual, thus the general ability that anybody can draw the same zig-zag line by saying "Zig-zag" would be lost.

7 Conclusion

We developed a novel drawing system, called Onomatopen, that enables users to draw various textured lines by saying aloud the corresponding onomatopoeia. In addition to drawing lines, users can edit and apply effects to images using onomatopoeia. This input technique enables users to draw in a more sensory-based manner and more enjoyable than with existing painting software. We will apply our interactive technique to other areas in which we can take advantage of onomatopoeia including animation, games and drawing maps.

References

1. *Grammar-based Continuous Speech Recognition Parser: Julius*. <http://julius.sourceforge.jp/>.
2. P. R. Cohen, M. Johnston, D. McGee, S. Oviatt, J. Pittman, I. Smith, L. Chen, and J. Clow. Quickset: multimodal interaction for simulation set-up and control. In *Proceedings of the fifth conference on Applied natural language processing*, pages 20–24, 1997.
3. S. Harada, J. A. Landay, J. Malkin, X. Li, and J. A. Bilmes. The vocal joystick: Evaluation of voice-based cursor control techniques. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 197–204, 2006.
4. S. Harada, T. S. Saponas, and J. A. Landay. Voicepen: Augmenting pen input with simultaneous non-linguistic vocalization. In *Proceedings of the 9th International Conference on Multimodal Interfaces*, pages 178–185, 2007.
5. T. Igarashi and J. F. Hughes. Voice as sound: Using non-verbal voice input for interactive control. In *Workshop on Interactive Systems and Software IX*, pages 7–10. Japan Society for Software Science and Technology, 2001.
6. L. Julia and C. Faure. Pattern recognition and beautification for a pen based interface. In *Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1)*, pages 58–63, 1995.
7. T. Komatsu and K. Hinata. Onomatopoeic filtering system for assisting user's intuitive expressions. In *The 22nd Annual Conference of the Japanese Society for Artificial Intelligence*, pages 1D2–10, 2008.
8. K. Kurihara, M. Gogo, J. Ogata, and T. Igarashi. Speechpen: Predictive handwriting based on ambient multimodal recognition. In *Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 851–860, 2006.
9. Y. Murakami. The inquiry of phonetic symbolism hypothesis : Analyses of onomatopoeias by the clustering of the introduced components no phonemes, sd technique, associated nouns, and associated verbs. *The Japanese journal of educational psychology*, 28(3):183–191, 1980.
10. S. Oviatt, R. Lunsford, and R. Coulston. Individual differences in multi-modal integration patterns: what are they and why do they exist? In *CHI'05*, pages 241–249, 2005.
11. I. Tamori and H. Kakehi. *Paradise of Onomatopoeia, Echoic word, Imitative word*. Keiso shobo, 1993.