

Mutual Recall Between Onomatopoeia and Motion Using Doll Play Corpus

Takuya Takahashi^{1,2} and Yasuyuki Sumi^{1(\boxtimes)}

¹ Future University Hakodate, Hakodate, Hokkaido 0418655, Japan t-takahashi@sumilab.org, sumi@acm.org ² Presently, with GREE, Inc., Tokyo, Japan

Abstract. Onomatopoeia is used to describe the state and degree of movement. Since onomatopoeia is a linguistic symbol, it is expected that many people will recall the same image corresponding to the onomatopoeia. Is this really true? In this paper, we focus on playing with dolls, collect onomatopoeia uttered during doll play, and construct a corpus of co-occurrence relations between onomatopoeia and doll movements. In order to verify the value of this corpus, we confirm the recall of movement from onomatopoeia and vice versa. As an application example based on this corpus, we also present a prototype of "onomatopoeia camera," a system that automatically adds onomatopoeia and comic-like expressions to subjects' movements.

Keywords: Onomatopoeia \cdot Motion corpus \cdot SAX \cdot Onomatopoeia camera

1 Introduction

Onomatopoeia is commonly used in our daily conversations. For example, we Japanese use an onomatopoeia "teku teku" to describe the state of walking motion and use another onomatopoeia "sasa" to describe faster movement.

Onomatopoeia is often thought as childish words or words just for extra decoration. Onomatopoeia, however, has attractive characteristics that bridge between linguistic symbols and unfixed concepts associated with bodily movements, mental status, and surrounding situations. We believe the deeper understanding of onomatopoeia leads us to mutual understanding among people having diffent culture, and between people and robots.

To date, many books and dictionaries focusing on onomatopoeia have been published. However, it is not enough to gather and discuss onomatopoeia as linguistic information to approach the essence of the characteristics of onomatopoeia. In this paper, we build an onomatopoeia corpus by collecting motion data of subjects as well as uttered onomatopoeia at the same time in order to enable more data-driven discussion of onomatopoeia. This paper also shows our attempt to build a system which mutually recalls between motion and onomatopoeia based on the corpus. It is practically difficult to collect onomatopoeia and surrounding data comprehensively in daily life. So, we chose doll play as a target where we collect data related on onomatopoeia because we can observe many onomatopoeia is uttered along with doll movements. We collected onomatopoeic utterances and corresponding doll motion during doll plays for building an onomatopoeia corpus.

This paper describes our attempt to collect doll's motion data recalled from various onomatopoeia. We build a corpus consisting of onomatopoeia and corresponding motion. We analyze the corpus on how motion associated with onomatopoeia converges or disperses depending on doll players. Secondly, we show the diversion of the corpus to recall suitable onomatopoeia from various motion measured in other domains such as human actions and robot behaviors. Lastly, we also show applications based on the corpus, e.g., a prototyped system called "onomatopoeia camera", that automatically adds onomatopoeia and comic-like expressions to a photo according to the motion of the subject in the photo.

2 Related Work

There has been attempts to evaluate onomatopoeia quantitatively [6,7]. In these work, onomatopoeia was represented by multiple vectors corresponding to human impressions. Another work presented a method that enables the estimation and depiction of onomatopoeia in computer generated animation based on physical parameters [2]. There were studies focus on the sounds recalled from onomatopoeia [8, 10, 11]. In the work utilizing the features of onomatopoeia, there was an attempt to propose a system that can present the user with a list of onomatopoeia specific to a restaurant they are interested in [5]. In another study, they produced an onomatopoeic learning support tool for Japanese learners [12].

Thus, analysis of onomatopoeia is often done by focusing on the impressions and sounds of words. We believe, however, that analysis of linguistic information alone is insufficient. Therefore, we construct a corpus consisting of onomatopoeia and motion. We also propose a system for recalling onomatopoeia using the corpus. We observe what kind of onomatopoeia the proposed system recalls from motion queries in various domains. We also analyze the relationship between onomatopoeia and motion.

3 Motion Corpus

3.1 Corpus Consisting of Onomatopoeia and Motion

In this work, we construct a corpus that enables mutual recall between onomatopoeia and doll motion. The configuration of the corpus is a collection of doll motion data expressed by participants for each onomatopoeia (Fig. 1). Each doll motion data is a position coordinate of three axes and a rotation value of three axes, and each value is obtained by extracting a motif in the doll movement.

This chapter describes the construction of the corpus and the clustering of doll motions in each onomatopoeia in the corpus.



Fig. 1. Motion corpus consisting of onomatopoeia and corresponding motion of doll.

3.2 Data Collection of Doll Motion

As a situation setting for constructing the motion corpus, we focused on doll play in which onomatopoeia is likely to be uttered. Participants were presented with onomatopoeic words as shown in Fig. 2 and asked to express the movements evoked by the onomatopoeic words with a doll. We collected doll movements and onomatopoeias as label data. The participants were 11 male university students. The onomatopoeia presented to participants was 21 words, and they expressed doll movements three times at each onomatopoeic word. As a result, a total of 693 doll movements were collected.

The doll movement is time-series data consisting of three dimensional position coordinates and rotation values. We measured doll movements using optical motion capture system (OptiTrack system with eight cameras). The doll movements are defined as the X-axis for the front direction of the doll, the Y-axis for the vertical direction, and the Z-axis for the horizontal direction. The rotation of these three axes is Roll, Yow, and Pitch.

We used 21 onomatopoeic words that we had previously observed in playing with dolls (Table 1). In the doll play, each of the three participants held one doll and were asked to play with it freely on the table, which was recorded on video. From the video recordings of the doll play, we extracted the onomatopoeia frequently uttered by participants. We also added some onomatopoeic words from a Japanese onomatopoeia dictionary that are similar to the onomatopoeia collected in the doll play.

3.3 Extracting Motion Motif Using SAX

The range of movement of the collected dolls varied widely from participant to participant. For example, some people exaggerated the doll's walking movement, while others expressed it in a modest manner. It is not possible to quantify the



Fig. 2. Collecting motion data of doll movement associated by given onomatopoeia.

Table 1. Onomatopoeia prepared to build our corpus.

onomatopoeia expressing walking and running						
Noro Noro	Ta Ta Ta	Teku Teku	Da Da Da	Toko Toko	Dota Dota	Suta Suta
onomatopoeia expressing jumping and flying						
Pyon	Pyoon	Byon	Byoon	Hyuun	Pyuun	Byuun
onomatopoeia expressing collision						
Doon	Doka	Dosun	Poyon	Kotsun	Gotsun	Zudon

intrinsic proximity between the doll movements generated by multiple participants if the physical quantities related to the doll movements are used as they are. Therefore, to be able to handle the qualitative proximity of each doll movement, we extracted the movement motifs from the doll movements and converted them into string representations.

We used Symbolic Aggregate Approximation (SAX) as a method to extract motion motifs from doll movements. SAX is one of the methods proposed by Lin et al. [9] for compressing time-series data, which converts the time-series data into character strings. Therefore, it is possible to use converted character strings to apply to pattern search and natural language processing for time-series data. In this study, the character string obtained by SAX is the automaton in one doll movement [1,3,4]. The procedure of SAX is shown below (Fig. 3).

- 1. Normalizing time-series data and divid into w equal sized frames. Mean value of the data falling within a frame is calculated.
- 2. Determining breakpoints that will produce equal sized areas under Gaussian curve. Set the character of each area.
- 3. Converting the mean value of each frame to a character.

The features of SAX include character string length(w) and character type(a) parameters after converting. We set the character string length to 20 and the



Fig. 3. String conversion of time-series data by SAX.

character kind 3 (w = 20, a = 3). We applied SAX to each of the three dimensional position coordinates and rotation values included in the obtained 693 doll movement data.

SAX converts data to relative values in order to normalize time-series data. We can compare with other data and measure the similarity between data, using the value converted to a character string at SAX. The threshold value for converting to a character string is automatically set to follow the normal distribution of the normalized time-series data. Therefore, SAX can mechanically convert unknown time-series data.

3.4 Cluster Analysis of Our Motion Corpus

We analyzed a hierarchical clustering of how the collected movements from doll play were grouped in each onomatopoeia in the motion corpus. Figure 4 on the left is a tree diagram obtained by hierarchical clustering, and the percentage of each class belonging to each onomatopoeia was calculated (right).

Most of the data for walking and running onomatopoeia were collected in C3. When these doll movements were checked, it was found that they were walking on the tabletop (Fig. 5). This indicates that the motions recalled from walking and running onomatopoeia are easy to unite independently of each other. From this, it was confirmed that the movements recalled by walking and running onomatopoeia are easily grouped into one class without depending on the person.

On the other hand, jumping and flying onomatopoeia tended to fall into the C1 and D classes, but were found to be dispersed in each class. The doll movements performed by the subject, which involved jumping and flying, were described by some as expressing a high mountain-shaped flight, while others as gliding horizontally with respect to the ground (Fig. 6). Therefore, it is considered that there was a difference in the vertical direction in the doll movement. From these results, it was confirmed that the motions recalled from the onomatopoeia that jumps and flies depend on people and belong to various classes.



Fig. 4. Cluster analysis of doll motion data (left figure). Class distribution of motion data of each onomatopoeia (right figure).



Fig. 5. An example of motion recalled from walking onomatopoeia.

Collision-type onomatopoeia tended to fall into the C1 and D classes but was found to be dispersed in each class. We confirmed the doll movements expressing the collision type onomatopoeia, various collision expressions were seen, such as colliding with something while walking or flying, falling down (Fig. 7). Therefore, it is considered that the difference between each motion data value was generated and distributed to each class. From these results, it was confirmed that the motions recalled from the collision type onomatopoeia tended to depend on each individual and belonged to each class.

4 Recalling Onomatopoeia from Motion

We prototyped a system that recalls onomatopoeia from arbitrary motion data. Figure 8 is an overview of the system.



Fig. 6. An example of motion recalled from jumping and flying onomatopoeia.



Fig. 7. An example of motion recalled from collision onomatopoeia.

4.1 Measuring Motion as Query

One of our goals is to recall onomatopoeia from movements in various domains. However, the range of motion as measured by the domain varies. For example, there are some ranges of motion such as doll play on a table and human physical movement. The information necessary for a motion query are three dimensional position coordinate and rotation. This section describes how to measure motion.

Doll movement in doll play or small robot motion are small range. For these domains, we used the motion capture system to measure the motion. We attached a few motion capture markers to a object of interest and created a rigid body. The position coordinate value and the rotation value of the created rigid body were measured. The motion capture system has high measurement accuracy of motion because it use many tracking camera. However, this system set up many cameras in a space. The measurable range of this system is small.



Fig. 8. Recalling onomatopoeia using our motion corpus.

The range of human activity related to walking and jumping movements is very large compared to playing with dolls. We used acceleration sensor and gyroscope sensor in this domain. We attached these sensors to a object and measured motion data. The position coordinate value and rotation value is obtained by integrating the obtained motion data. The measurable range using these sensors is big because these are easily mounted. However, the motion capture system has low measurement accuracy of motion.

4.2 Calculating Similarity Among Motions

The system performs a similarity search using motion data as the query. We used the Levenshtein distance as the method for similarity search. Levenshtein distance is a measure of the similarity between two strings. The distance is the number of deletions, insertions, and substitutions required to transform query characters into purpose characters. The greater the Levenshtein distance, the more different the characters are. This system calculate the Levenshtein distance between motifs of motion query extracted by SAX and each motion data in motion corpus. After performing the similarity search based on the motion, onomatopoeia of motion data with a shorter similarity distance is taken as the searching result.

5 Applying Doll Corpus into Other Domains

We observed that the proposed system recalls what kind of onomatopoeia from motion queries of various domains. We focused on doll play, small robot and human body movement in daily life as a domain. We confirm if motion dictionary have versatility through recalling onomatopoeia from multiple domains.

5.1 Within "Doll" Domain

Measurement of Doll Play. We observed that proposed system recalls what kind of onomatopoeia from doll motion in doll play (Fig. 9). Motion with high degree of freedom than motion corpus doll motion can be seen in doll play. In the case of same as motion corpus domain, we confirm that whatever proposed system can recall an onomatopoeia suitable for motion.

Three participants hold a doll and play with dolls on the table about five minutes. We prepared three types of dolls: hero, monster, and anthropomorphic tower. A projector displayed a wrestling stage on the table. We asked three participants to play with dolls freely. We measured the dolls motion using motion capture system.



Fig. 9. Doll play: Playing with dolls on a table with a wrestling stage projected on it.

Recalling Onomatopoeia. As a result of onomatopoeia recalling from doll motion in doll play, walking motion recalled walking system onomatopoeia such as "*Suta Suta*" and "*Toko Toko*". Collision and overturning motion recalled collision system onomatopoeia such as "*kotsun*" and "*gotsun*". Motion queries in doll domain recalled onomatopoeia of a system suitable for motion.

5.2 Applying to "Robot" Domain

Measurement of Robot Motion. We observed that proposed system recalls what kind of onomatopoeia from robot motion (Fig. 10). We confirm that whether motion dictionary have versatility for motions other than the doll.

We used Jumping Sumo which is a toy drone as a robot. Jumping Sumo is a two wheels drone. This robot can be controlled by smartphone application. This drone has preset of motion and it can motion such as jumping and turning. We measured the preset motions using motion capture system.

Recalling Onomatopoeia. As a result of onomatopoeia recalling from robot motion, robot's rectilinear motion recalled walking system onomatopoeia such



Fig. 10. Measuring robot motion.

as "*Ta Ta Ta*" and "*Toko Toko*". Jumping motion recalled jumping and flying system onomatopoeia such as "*Byoon*" and "*Hyuun*". Robot domain motion recalled onomatopoeia of a system suitable for motion.

5.3 Applying to "Human" Domain

Measurement of Human Body Movements. We observed that motion dictionary recalled what kind of onomatopoeia from human movement (Fig. 11). People can express the more exaggerated motion more than the body movement by using dolls. For example, by using doll we can express jumping higher and flying motion that human cannot do. However, we focus on motif of motion and compare these motions. In the motif of motions, walking and jumping are similar. Therefore, we assume that human body movement recalls onomatopoeia suitable for human movements.

We collected motion data that three participants attach the smartphone on the chest. We measured body movements using acceleration sensor and angular velocity sensor are mounted on a smartphone. We asked three participants to action three motions such as walking, running and skipping on a certain distance.



Fig. 11. Measuring human motion.

Recalling Onomatopoeia. As a result of onomatopoeia recalled from body movements in daily life, walking motion like walking and running recalled collision system onomatopoeia such as "*Kotsun*" and "*Doka*". Jumping motion like skipping recalled flying system onomatopoeia such as "*Byuun*" and "*Hyuun*". In the case of recalling onomatopoeia from human domain motion queries, some motion queries recalled onomatopoeia which are not suitable for system.

6 Discussion on Onomatopoeia Recalling

We observed that proposed system recalls what kind of onomatopoeia from motion queries of various domains using motion corpus. As a result, some motions such as doll, robot and human recalled suitable onomatopoeia. Jumping motion queries of these domains recalled jumping system onomatopoeia. We checked motif of the motion expressing jumping system onomatopoeia in motion corpus. The motif of these motion were characterized in the vertical direction (Table.2 and 3). These motif of vertical direction included several kinds of characters. It is conceivable that jumping and flying system onomatopoeia is likely to be recalled in jumping and flying motions because these motions are seen in the vertical direction motif.

Table 2. An example of position motif	obtained from a robot's jumping motion.
---------------------------------------	---

Position	Motif
Х	aaaaaaabbbbbbccccccc
Y	aaaaabcccccccbbaaaa
Z	bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb

Table 3. An example of position motif obtained from a	a person's skipping motion.
---	-----------------------------

Position	Motif
Х	aaaaabbbbbbbbbbbbbbccccc
Y	aaabccbaacccbabccbaa
Z	bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb

We confirmed that the onomatopoeia of the collision system was recalled against the motion seen in the doll play. Participants were attacking and treading other dolls with his dolls. These motions recalled collision system onomatopoeia. These doll motions rotate in the middle of walking and flight operations, often colliding with the other dolls. The motif of these motion were characterized in the pitch rotation (Table 4). Collision system onomatopoeias are easy to be recalled for motion including pitch rotation. In one scene, walking motion of monster type doll recalled the collision system onomatopoeia. The monster type doll was attacked by another doll while walking in this scene. It was also seen that onomatopoeia recall combined with relationship with other doll motion was done. It was seen that there was difference in the motion motif representing the rotation values.

Table 4. An example of rotation motif obtained from a doll's collision motion.

Rotation	Motif
Roll	bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
Yaw	bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
Pitch	bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb

On the other hand, walking motion of human could not recall walking system onomatopoeia. We compared the walking motion motif of human and doll. Humans move forward while maintaining their posture. The dolls were advancing while swaying alternately between left and right (Table 5). Also, There was a pattern that the doll was advancing while moving slightly up and down. It was seen from each motifs of walking motion that this doll motion affects the value of rotation value. People can express the more exaggerated or impossible motion more than the human motion by using dolls. For example, flying motions and collision motions. In onomatopoeia recall, the motion motif of three dimensional position and rotation were used for similarity search. However, It is conceivable that similarity search using all value quantities is not desirable because human do not much rotational motion. When recalling onomatopoeia in the human domain, we have to examine feature quantities used for similarity search.

 Table 5. Comparing Yaw rotation motifs between walking motions of human and doll.

Motion actor	Yaw rotation motif
Human	bbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
Doll	bbbcbcaccbbbbbbbaaaab

7 Application Systems

7.1 Onomatopoeia Camera

We created an onomatopoeia camera that adds onomatopoeia following the movement of the subject in the camera and performs comic-like effects on photographs. A smartphone is attached to the subject's chest, and the acceleration value is acquired from the subject's smartphone by the shutter timing of the camera app created (Fig. 12). This application recall onomatopoeia displayed on the camera application using the motion corpus constructed from the obtained acceleration values (Fig. 13).



Fig. 12. Usage example of onomatopoeia camera.

We had several people experience this camera application and observed it. App users enjoyed seeing the onomatopoeia appearing in the first photo taken. Gradually, they thought about the patterns between onomatopoeic words and movements of humans and tried and errored what kind of onomatopoeia would be displayed. We think that users can learn cognitive science, such as what kind of onomatopoeia is recalled from human body movements, through camera applications.

This application can also manually input effect lines such as concentration lines, which are a type of manga expression, as shown in the photo on the right of the figure. By applying such comics effects to the photographed pictures, it is thought that the state at the time of photography can be conveyed clearly and expressively. Also, comics culture in Japan has been evaluated abroad, and we believe that the created applications will be useful for intercultural communication and onomatopoeic learning for non-native Japanese speakers.

7.2 Augmented Doll Play

This application measures a doll motion in doll play and recalls onomatopoeia from doll motion. The recalled onomatopoeia is displayed on the table (Fig. 14). We observed the situation of the players who played this AR system. The player tried what kind of onomatopoeia would be recalled for a certain motion. Other players looked at the onomatopoeia displayed on the table and moved



Fig. 13. Examples of onomatopoeia camera: The photo on the center has comic-like effect which is automatically added associated with the selected onomatopoeia. The photo on the right shows a monochrome effect.

the dolls. While the system recalled onomatopoeia from motion, the players recalled motion from onomatopoeia. We can learn motion and language from onomatopoeia. With this AR system, this language education for children or expansion of physical expression communication can be augmented.



Fig. 14. Projection of recalled onomatopoeia during doll play.

8 Conclusions

This paper showed our attempt to build a corpus consisting of onomatopoeia and corresponding motion of subjects. We built our corpus by choosing doll play as a target since we could observe various onomatopoeia.

We proposed a system which recalls onomatopoeia from doll's motion. We extracted motif from doll's motion data in order to abstractly represent the

motion data associated with corresponding onomatopoeia. We confirmed if our system could successfully recalled corresponding onomatopoeia from motion queries across various domains. The result shows that the motion corpus associated with jumping and collision onomatopoeia of doll could be successfully diverted to both robot and human domains. Human walking motion, however, could not recall appropriate onomatopoeia associated with walking because the rotation values of them are quite different.

As application examples based on this corpus, we presented a prototype of "onomatopoeia camera", a system that automatically adds onomatopoeia and comic-like expressions to subjects' movements, and augmented doll play.

References

- Canelas, A., Neves, R., Horta, N.: Multi-dimensional pattern discovery in financial time series using SAX-GA with extended robustness. In: Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation (GECCO '13 Companion), pp. 179–180. ACM, NY, USA (2013). http://doi.acm. org/10.1145/2464576.2464664
- Fukusato, T., Morishima, S.: Automatic depiction of onomatopoeia in animation considering physical phenomena. In: Proceedings of the Seventh International Conference on Motion in Games (MIG '14), pp. 161–169. ACM, NY, USA (2014). http://doi.acm.org/10.1145/2668064.2668096
- Garg, Y., Poccia, S.R.: On the effectiveness of distance measures for similarity search in multi-variate sensory data: effectiveness of distance measures for similarity search. In: Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval (ICMR '17), pp. 489–493. ACM, NY, USA (2017). http:// doi.acm.org/10.1145/3078971.3079009
- Gomes, E.F., Jorge, A.M., Azevedo, P.J.: Classifying heart sounds using multiresolution time series motifs: An exploratory study. In: Proceedings of the International C* Conference on Computer Science and Software Engineering (C3S2E '13), pp. 23–30. ACM, NY, USA (2013). http://doi.acm.org/10.1145/2494444.2494458
- Kato, A., Fukazawa, Y., Sato, T., Mori, T.: Extraction of onomatopoeia used for foods from food reviews and its application to restaurant search. In: Proceedings of the 21st International Conference on World Wide Web (WWW '12 Companion), pp. 719–728. ACM, New York, NY, USA (2012). http://doi.acm.org/10.1145/ 2187980.2188192
- Komatsu, T.: Quantifying Japanese Onomatopoeias: toward augmenting creative activities with onomatopoeias. In: Proceedings of the 3rd Augmented Human International Conference (AH '12), pp. 15:1–15:4. ACM, NY, USA (2012). http://doi. acm.org/10.1145/2160125.2160140
- Komatsu, T.: Choreographing robot behaviors by means of Japanese onomatopoeias. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI'15) Extended Abstracts, pp. 23–24. ACM, NY, USA (2015). http://doi.acm.org/10.1145/2701973.2702020
- Kwon, S., Kim, L.H.: Sound sketching via voice. In: Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication (ICUIMC '11), pp. 48:1–48:4. ACM, New York, NY, USA (2011). http://doi.acm. org/10.1145/1968613.1968672

- Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A symbolic representation of time series, with implications for streaming algorithms. In: Proceedings of the 8th ACM SIG-MOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD '03), pp. 2–11. ACM, New York, NY, USA (2003). http://doi.acm.org/ 10.1145/882082.882086
- Miyazaki, A., Tomimatsu, K.: Onomato Planets: physical computing of Japanese onomatopoeia. In: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction (TEI '09), pp. 301–304. ACM, NY, USA (2009). http://doi. acm.org/10.1145/1517664.1517726
- Oh, J., Kim, G.J.: Effect of accompanying onomatopoeia with sound feedback toward presence and user experience in virtual reality. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18), pp. 81:1–81:2. ACM, New York, NY, USA (2018). http://doi.acm.org/10.1145/ 3281505.3283401
- Yusuf, M., Asaga, C., Watanabe, C.: Onomatopetal: developing a Japanese onomatopoeia learning-support system utilizing native speakers cooperation. In: Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT '08) - Volume 03, pp. 173–177. IEEE Computer Society, Washington, DC, USA (2008). https://doi.org/10.1109/ WIIAT.2008.300