A Presentation Avatar for Self-Review

Keisuke INAZAWA* & Akihiro KASHIHARA

The University of Electro-Communications, Japan 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585 *keisuke.inazawa@uec.ac.jp

Abstract: Presentation is one of the most important activities for researchers. For proper presentation, self-rehearsal is often conducted. In self-rehearsal, presenters usually make presentation to themselves with their PC. However, they often miss finding some points to be modified, since they need to review it concurrently with making their presentation. On the other hand, there is another way for self-reviewing, in which they could make a video of their presentation and then check it out. Although it allows them to direct more efforts to review, they would feel quite uncomfortable due to their looks and voice on the video. Our approach to this issue is to design presentation avatar, which allow the presenters to objectively self-review their own presentation avatar, which reproduces presentation the presenters make. This paper also reports a case study using it, in which we compare self-review with presentation video and self-review with presentation avatar. The results suggest that presentation avatar has a potential to promote self-reviewing to improve presentation.

Keywords: Presentation, Avatar, Self-review, Rehearsal, Objective review

1. Introduction

Presentation is one of the most important activities for researchers. They are accordingly required to develop skill in presenting their work properly. Proper presentation needs several rehearsals, which involves reviewing the presented contents, oral explanation, gesture, etc. to improve the presentation. It is crucial particularly for unskillful researchers to conduct presentation rehearsal. Such rehearsal would also contribute to developing presentation skill.

In research laboratory, presentation rehearsal is often conducted in two ways, which are rehearsal with peers (lab members) and self-rehearsal. In rehearsal with peers, presenters could receive peer reviews indicating points to be modified (Ryo Okamoto, and Akihiro Kashihara, 2007). In self-rehearsal, they need to make presentation and review it by themselves from an audience point of view to find the modified points. Self-rehearsal is often conducted before rehearsal with peers.

The main issue addressed in this paper is how to allow unskillful presenters as learners to self-review and improve their presentation. In self-rehearsal, presenters usually make presentation to themselves with their PC. However, they often miss finding some points to be modified, since they need to review it concurrently with making their presentation. On the other hand, there is another way for self-reviewing, in which they could make a video of their presentation and then check it out. Although it allows them to direct more efforts to review, they would feel quite uncomfortable due to their looks and voice on the video. As Holzman et al. pointed out (Holzman, Philip S., and Clyde Rousey, 1966), such uncomfortable sense often occurs from discrepancy between the voice one hears in one's head and the recorded voice. In the same way, the presenters could feel uncomfortable with discrepancy between the looks one expects and the recorded looks. This uncomfortableness prevents the presenters from objectively self-reviewing their recorded presentation from audience viewpoints.

In order to address this issue, we are designing presentation avatar (P-Avatar for short), which acts as computer agent or educational robot for reproducing the presentation presenters make, and for making presentation as their proxy. The main aim of P-Avatar is to allow the presenters to concentrate on self-review of their own presentation without any uncomfortable sense and to find points to be modified.

In this paper, we propose and demonstrate P-Avatar that reproduces presentation learners make. This paper also reports a case study using it, in which we compare self-review with presentation video and self-review with P-Avatar. The results suggest that P-Avatar has a potential to promote self-reviewing to improve presentation.

2. Self-Review

2.1 Presentation Rehearsal

Presentation rehearsal allows learners to gain reviews including points to be modified and to improve their presentation. It also gives them an opportunity to develop their presentation skill. It can be modeled as a cyclic process involving three phases as shown in Figure 1, which are preliminary presentation, review, and modification. In the phase of preliminary presentation, learners rehearse presentation with a presentation document (P-document for short) such as PowerPoint/Keynote one according to the context expected. In the review phase, they check out their presentation with peers including more skillful lab members or by themselves to gain points to be modified. In the phase of modification, the learners are expected to follow the review results to modify the P-document, oral explanation, gesture, etc. By repeating these phases, the learners can improve their presentation before actual presentation.

In developing presentation skill, it is indispensable for unskillful learners to enhance the ability

to self-review (Nancy, D., 2008), although peer review is an instructive way for improving presentation. In self-review, learners are required to have an audience view to objectively check how they used P-document, how they made oral explanation and non-verbal behavior such as gesture and facial expression, and also to check the contents of P-document and oral explanation, etc.

Such objective self-review allows them to find out what should be modified. The learners are particularly expected to become aware of excess/deficiency/suitability in the contents, design, and slide order of the P-document. They are also expected to become aware of improper non-verbal behavior including eye contact to the audience, behavioral habits, pointing to the slide, and improper oral explanation including speed, emphasis on keywords/sentences, connection between slides, intonation, intervals, haste, etc.

However, it is not so easy for unskillful learners to gain such awareness of points to be modified from self-review. How to promote self-reviewing to increase awareness of the modified points in presentation is the main issue addressed in this paper.

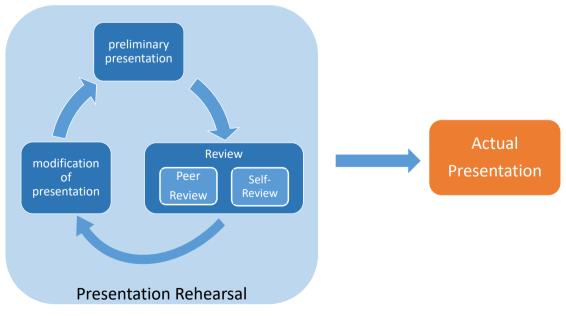


Figure 1. Rehearsal model.

2.2 Problems in Self-Review

The normal way for learners to do presentation rehearsal involving self-review is to make presentation to themselves with their PC as shown in Figure 2(a). However, it is quite difficult to self-review it in detail since the self-review process is conducted concurrently with presentation. Some points to be modified could be accordingly missed.

On the other hand, there is another way of the self-rehearsal as shown in Figure 2(b), which is to make a video of their presentation and then to check it out. This allows the learners to direct more efforts to review. But, they would feel quite uncomfortable due to their looks and voice on the video. Such uncomfortable sense would occur from discrepancy between the looks/voice one expects in one's head and the recorded looks/voice. It is quite difficult for the learners to review the presentation objectively with this uncomfortable sense, which prevents them from concentrating on self-review. Even though they could overcome the uncomfortable sense, in addition, there would be a limit to finding points to be modified. Although objective perspective contributes to gaining awareness of points to be modified, it is quite difficult to hold it in self-review.

In order to resolve the above problems, it is necessary to remove the uncomfortable sense to be brought about in self-review to promote checking out the presentation and enhance awareness of the modified points.

2.3 Related Work

Related work on supporting self-rehearsal of presentation has been mainly addressing how to automatically analyze non-verbal behavior and oral explanation from motion capture of presentation. For instance, Kurihara et al. has proposed a system, which analyzes degree of eye contact to audience, oral speed, intonation, and hesitation, from captured motion and then to present the analyzed statistical data to learners as review results during/after their presentation (Kazutaka Kurihara, et al., 2007). The learners could have training of self-rehearsal with the system, and could become aware of non-verbal behavior and oral explanation to be modified with such data. Although such analyzed data seems instructive for presentation improvement, self-review is limited within the non-verbal behavior defined in advance by the system. The system does not provide learners with any support that promotes reviewing their P-documents.

Zhao et al. has also proposed a system, which detects and evaluates non-verbal behavior necessary for proper presentation such as posture, attitude, gesture, etc (Xinbo Zhao, Takaya Yuizono, and Jun Munemori, 2015). If the system detects improper behavior during presentation, learners immediately receive feedback with evaluated information. After presentation, they can also check out the presentation video with the evaluated information. However, they would still have difficulties in gaining awareness of points to be modified since the system provides no support for decreasing uncomfortable sense brought about by looking their looks and listening to their voice in the self-review process.

Chollet et al. has proposed a system using virtual audience (Chollet et al., 2015). The system analyzes eye contact to audience and number of pause fillers with audiovisual sensors, and then presents either directly visualized performance measures or implicit nonverbal feedback from the virtual audience during their presentation. The virtual audience will behave positively when the presenter is

performing well, and behave negatively when the presenter is not performing well. For example, they smile when the presenter make eye contact with them, or they look away when a pause filler occurs. In this system, however, it is also difficult to gain awareness of points to be modified in detail without support after presentation.

P-Avatar proposed in this paper does not intend to generate review results as feedback, but intends to allow learners to review their presentation in a self-directed and objective way.



(a) Presentation with PC



(b) Presentation with video-recording <u>Figure 2</u>. Self-rehearsal of presentation.

3. Self-Review with Presentation Avatar

3.1 What is Presentation Avatar?

In order to promote self-directed and objective self-review of presentation, we design P-Avatar. The main purpose of P-Avatar is to remove uncomfortable sense brought about in checking out presentation video, which occurs from discrepancy between expected looks/voice and recorded ones. We are currently considering P-Avatar a virtual agent running on computer and an educational robot operating in real world. In this paper, we introduce P-Avatar as a virtual agent.

P-Avatar has two main roles, which are to reproduce presentation learners make, and to produce presentation as their proxy. In presentation reproduction, learners first make preliminary presentation. P-Avatar second records it and captures the slides used, timing of slide change in using their P-document, and their voice/non-verbal behavior. P-Avatar third reproduces the presentation with the slide

information as exactly as possible with the captured data on computer screen. But, the voice is changed with P-Avatar's one. The learners then check out the presentation reproduced by P-Avatar. In this way, the presentation reproduction allows the learners to objectively self-review their presentation to gain awareness of points to be modified. In order to further enhance such awareness, we are now considering an effective reproduction, in which P-Avatar exaggerates some behaviors different from average ones or some behaviors essential to presentation. Such exaggerated reproduction is out of our scope in this paper.

In proxy presentation, learners do not need to make preliminary presentation, but prepare their Pdocument, its oral manuscript, timing information for changing the slides in using the P-document, and information about pointing to keywords/sentences as their gesture in advance. P-Avatar uses the prepared data to produce the presentation instead of the learners. They then check out the produced presentation by P-Avatar for presentation improvement. In this way, the proxy presentation allows learners to review presentation expected from the prepared data without making their preliminary presentation. In particular, it would be useful for improving P-document and reconsidering how to use it.

In this paper, we focus on how to promote self-review of presentation made by learners. In the following, let us accordingly demonstrate P-Avatar with the role of presentation reproduction.

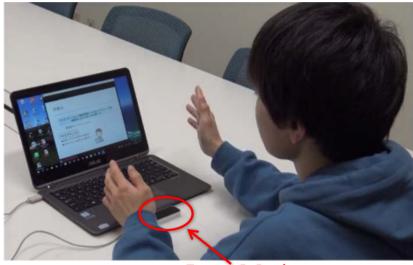
3.2 P-Avatar System

We have developed a P-Avatar system using Unity Technologies (Unity - Game engine, tools and multiplatform, 2017), in which P-Avatar is implemented as virtual agent on computer. Using Leap Motion (Leap Motion, 2017) and microphone as input devices, the system records and captures presentation learners make as shown in Figure 3 (a). In reproducing the presentation with the captured data, the current system uses the secondary creation of the character "Hatsune Miku" (About HATSUNE MIKU | CRYPTON FUTURE MEDIA, n.d.) of Crypton Future Media, INC as P-Avatar as shown in Figure 3 (a), which is created from the model data "Tda formula Hatsune Miku Append" provided by Tda (Tda formula Hatsune Miku Append Ver1.10 – BowlRoll, 2017). This is created under Piapro Character License (piapro.net, n.d.). The system also integrates MMD 4 Mechanism (Stereoarts Homepage, 2017) created by Nora to make use of the model data on Unity.

Let us demonstrate how to reproduce the presentation with P-Avatar in the following. The system first converts a PPT file to image file, and displays the first slide as image in the user interface. Learners can change the slide into the next/previous one via keyboard. Animation embedded in the slide is turned off. Before presentation, learners can select the use of a pointer.

When learners start self-rehearsal, the system makes P-Avatar invisible, and induces them to pay more attention to their presentation. Instead of P-Avatar, the system displays a marker that indicates the position of the right hand of P-Avatar. If the pointer is used, this marker is not displayed. Their hand gesture is projected as movement of the marker or pointer in real time. The pointer/marker is displayed all the time during presentation. It allows the learners to know to which position/direction their right hand points. When they also change the slide to the next or previous one via keyboard, the system captures the timing. Leap Motion captures their hand gesture that includes pointing to the slide. The microphone records their oral explanation. Their utterance is not projected in real time. The system currently captures only hand gesture, but it seems most important in presentation. That is the main reason we focus on it as presentation gesture.

When the learners finish presentation, the system generates video with the recorded and captured data, in which P-Avatar reproduces the presentation. As for oral explanation, the system transforms the voice quality by adjusting the value of fundamental frequency and formant frequency. The fundamental frequency indicates voice height. The formant frequency also indicates resonance determined by the shape of a vocal tract or an oral cavity, and it characterizes individual voices. When they replay the video, the hand gesture and utterance captured are projected as hand and mouth movement of P-Avatar as shown in Figure 3(b). If the position of the pointer moves to the left of the screen beyond a certain line, P-Avatar faces toward the slide. The presentation slides are displayed and changed in the recorded timing. Figure 3(c) shows an example of presentation reproduction with P-Avatar, in which P-Avatar points to the slide with the pointer to explain.



(a) Presentation with P-Avatar system



Reproduction with F-Avatar

Figure 3. P-Avatar system.

4. Case Study

4.1 Preparation and Procedure

We had a case study with the purpose of ascertaining whether P-Avatar could promote self-reviewing presentation without uncomfortable sense. Compared to self-review with presentation video, we analyzed how many points to be modified could be found. We also prepared a questionnaire that asked which self-review enhanced awareness of modified points, concentration on review, and uncomfortable sense.

The participants were 9 graduate and undergraduate students in informatics and engineering. We set two conditions: (a) self-review with presentation video (Video condition), and (b) self-review with the P-Avatar system (P-Avatar condition). As within-subject design, each participant conducted self-review twice under these two conditions. The participants were divided into two groups called group V-PA and group PA-V. We randomly assigned 5 participants to group V-PA and 4 participants to group PA-V. Group V-PA first self-reviewed under the Video condition and then self-reviewed under the P-Avatar condition. Group PA-V self-reviewed in the reversed order of the two conditions.

This study included 2 sessions referred as Session I (presentation) and Session II (self-review). Before Session I, all participants were required to prepare their PPT document including the contents

of their research so that they can make presentation within 15 minutes. In Session I, they were required to carry out preliminary presentation twice with their prepared P-document. They were first required to present with their P-document projected on a screen in front of video camera. The presentation was video-recorded. They were then required to present with the P-Avatar system. They were also informed that they would be required to self-review their presentation two days later.

In Session II conducted after two days, all participants were required to check out their presentation twice with the recorded video and the P-Avatar system, and to find points to be modified. In this session, group V-PA was required to first watch the recorded video and then to use the P-Avatar. Group PA-V was also required to first use the P-Avatar system and then to watch the recorded video. We had an interval of 3 hours between two self-reviews. Before using the P-Avatar system, all participants were given an explanation about how to use it. The use of a pointer in P-Avatar was left up to them. Since P-Avatar was a female virtual character, in this study, the system transformed their recorded voice by raising the fundamental frequency and formant frequency value. In each self-review, we gave all participants a handout of their PPT document. The participants were required to annotate the corresponding slide in the handout with the points to be modified. After Session II, all participants were required to answer a questionnaire, in which they were asked which self-review had more influence on awareness of modified points, concentration on self-review, and uncomfortableness.

In order to ascertain whether P-Avatar could contribute to finding points to be modified in presentation, we compared the numbers of modified points found in Session II, which we summed up from annotations made in the handouts. We also ascertained the potential of P-Avatar for promoting self-review of presentation from the answers to the questionnaire. The hypotheses we set up in this study were as follows:

- H1: P-Avatar allows finding more points to be modified, and
- H2: P-Avatar provides more awareness of modified points, concentration on self-review, and less uncomfortable sense.

4.2 Results

Table 1 shows the number of points to be modified that were obtained in self-review by the participants. However, this table includes modified points related to slide animation and body gesture in Video condition, which could not be obtained in Self-Avatar condition. It also includes modified points related to operations of P-Avatar, which could not be obtained in Video condition. We accordingly removed the number of such modified points from Table 1 to make Table 2. Although only 3 participants obtained more points to be modified with P-Avatar in Table 2, 7 of 9 participants obtained more or equal modified points with P-Avatar than recorded video in Table 2. In particular, all participants in group PA-V obtained more modified points per slide in Self-Avatar condition were higher than Video condition in Table 2. These results suggest that hypothesis H1 is supported.

Following the results in Table 2, we further divided the points to be modified into ones for Pdocument, ones for gesture, and ones for oral as shown in Table 3. Table 4 shows the average numbers of points to be modified for P-document, gesture, and oral obtained in both conditions and both groups. Overall, there were fewer points to be modified for gesture. Except modified points for P-document and oral in group V-PA, the average numbers of modified points per slide obtained in P-Avatar condition were higher than the ones in Video condition. From the result of the one-sided t-test with logarithmic transformation, there was a tendency of significant difference between the average numbers of modified points for P-document in Video condition and P-Avatar condition (t $_{(8)} = 0.0675$, p <. 10).

Table 1. Number o			Video condition		P-Avatar condition	
Group	Participant	Number of slides	Modified points	Average points per slide	Modified points	Average points per slide
	А	28	35	1.250	24	0.857
0	В	35	37	1.057	28	0.800
Group V-PA	С	40	17	0.425	9	0.225
V-FA	D	25	12	0.480	14	0.560
	E	23	18	0.783	17	0.739
	F	30	8	0.267	9	0.300
Group	G	32	6	0.188	12	0.375
PA-V	Н	40	20	0.500	19	0.475
	I	35	16	0.457	13	0.371
Average points per slide in group V-PA			0.799		0.636	
Average points per slide in group PA-V		0.353		0.380		
Average points per slide in both groups		0.601		0.523		

Table 1: Number of modified points found in self-review.

Table 2: Number of modified points except the ones removed.

		Video condition		P-Avatar condition		
Group	Participant	Number of slides	Modified points	Average points per slide	Modified points	Average points per slide
	A	28	24	0.857	24	0.857
0	В	35	30	0.857	28	0.800
Group V-PA	С	40	13	0.325	9	0.225
	D	25	8	0.320	14	0.560
	E	23	15	0.652	17	0.739
	F	30	8	0.267	9	0.300
Group	G	32	5	0.156	12	0.375
PA-V	Н	40	17	0.425	18	0.450
	I	35	12	0.343	13	0.371
Average points per slide in group V- PA			0.602		0.636	
Average points per slide in group PA- V			0.298		0.374	
Average points per slide in both groups			0.467		0.520	

Table 3: Number of points to be modified for P-document, gesture, and oral.

		Video condition			P-Avatar condition		
		P- document	Gesture	Oral	P- document	Gesture	Oral
Group	Participant	Total	Total	Total	Total	Total	Total
	А	10	4	10	7	6	11
Group V-PA	В	11	7	12	20	0	8
	С	1	0	12	3	0	6
	D	0	0	8	5	0	9
	E	10	1	4	9	3	5
Group PA-V	F	1	0	7	3	1	5
	G	0	0	5	1	0	11
	Н	3	0	14	1	0	17
		4	0	8	5	0	8

	Video condition			P-Avatar condition		
	Р	G	0	Р	G	0
	Average points per slide					
group V-PA	0.226	0.077	0.299	0.298	0.069	0.270
group PA-V	0.056	0.000	0.242	0.075	0.008	0.291
both groups	0.150	0.043	0.274	0.199†	0.042	0.279

Table 4: Averages of points to be modified per slide.

† One-sided t-test: p <.10

Table 5 next shows the results of the questionnaire, which consisted of four questions. From the results of accurate binomial test for Q1 to Q3, there were significant differences between recorded video and P-Avatar (p = 0.0898, p <. 10 for Q2; and p = 0.0020, p <.01 for Q3). The answers to Q1 and Q2 suggest the potential of P-Avatar for promoting awareness of points to be modified and concentration on self-review. The answer to Q3 also suggests that P-Avatar has a potential to remove uncomfortable sense occurring in self-review. These results support H2.

Table 5: Questionnaire and results.

Questions	Number of participants		
Q1. Which self-review promoted	Recorded video	P-Avatar	
obtaining points to be modified?	3	6	
Q2. Which was better for	Recorded video	P-Avatar	
concentration on review?	2	7†	
Q3. Which made you feel	Recorded video	P-Avatar	
uncomfortable?	9	0**	
Q4. Was there any inconvenience	Yes	No	
in using the P-Avatar system?	7	2	
	*	* p<.01, † p <.10	

4.3 Discussion

Table 2 and Table 4 suggest that P-Avatar contributes to finding more points to be modified for presentation particularly for P-document. Table 5 also suggests that P-Avatar enhances awareness of modified points, and concentration on self-review, and decreases uncomfortable sense. From these results, we can say that P-Avatar promotes self-review of presentation.

In this case study, self-review was conducted twice. The second self-review could accordingly obtain more points to be modified than the first review since it is easier to become aware of modified points obtained in the first review. The number of modified points in the first review tends to become lower than the one in the second review. As shown in Table 2, however, all participants in group PA-V obtained more modified points with the P-Avatar system even in the first review. This also suggests that P-Avatar strongly enhance awareness of points to be modified.

In Table 4, there were no significant differences about gesture and oral. As for gesture, five participants did not find any modified points. This suggests that they were able to perform expected gestures sufficiently or they did not regard gestures as modified points. As for oral, three participants pointed out in the questionnaire that the volume of presentation reproduced by P-Avatar was small, and that the sound had a little bit of noise. We need to address these problems as our future work.

As for Q1 and Q2 in the questionnaire shown in Table 5, there were 3 participants who selected recorded video. However, all the participants obtained more points to be modified with P-Avatar. They commented on shortcomings of the P-Avatar system such as no slide animation, difficulty in projecting hand gesture on P-Avatar, low volume of P-Avatar, P-Avatar's looks, etc., which seem the reasons they

did not select P-Avatar in Q1 and Q2. We need to correct these shortcomings to refine the P-Avatar system, which is our future work.

From the result of Q4 in the questionnaire shown in Table 4, in addition, most of the participants felt inconvenient to use the P-Avatar system. In this case study, there were also points to be modified, which could be found with recorded video and could not be found with the P-Avatar system. The representative points are slide animation, and non-verbal behavior (except hand gestures) such as body movement, gaze, facial expression, etc. In future, we accordingly need to redesign the use of the P-Avatar system, and to address how to reproduce slide animation and how to capture and reproduce the non-verbal behavior to improve the potential of self-review with P-Avatar.

5. Conclusions

In this paper, we have proposed P-Avatar, which promotes self-directed and objective self-review of presentation. The P-Avatar system allows learners to remove uncomfortable sense, which would occurs from their looks and voice on the video. It also allows them to objectively self-review their presentation to gain awareness of points to be modified. The results of the case study with the P-Avatar system suggested that P-Avatar has a potential to promote self-reviewing to improve presentation.

In future, we will refine the P-Avatar system so that it can particularly reproduce slide animation and more adequate motion capture. In addition, we will also develop a P-Avatar system, which carries out proxy presentation or exaggerated reproduction of presentation, and which uses educational robot as P-Avatar.

References

- Ryo Okamoto, Akihiro Kashihara. (2007). Construction of Presentation Review Environment with Realtimecreated Hyper Video Technique: *IEICE Technical Report Order Information*, Educational Technology Vol. 106, No. 583, pp133-138.
- Holzman, Philip S., and Clyde Rousey. (1966). The voice as a percept: *Journal of Personality and Social Psychology*, Vol. 4(1), pp79-86.
- Nancy, D. (2008). Slide:Ology: The Art and Science of Creating Great Presentations. USA: Oreilly & Associates Inc.
- Kazutaka Kurihara, Masataka Goto, Jun Ogata, Yosuke Matsusaka, and Takeo Igarashi. (2007). Presentation sensei: a presentation training system using speech and image processing: ICMI '07 *Proceedings of the 9th international conference on Multimodal interfaces*, pp358-365.
- Xinbo Zhao, Takaya Yuizono and Jun Munemori. (2015). Development of A Presentation Support System for Nonverbal Expression Skills: *IPSJ SIG Notes 2015-Groupware and Network Services 6*, pp1-6.
- Mathieu Chollet, Torsten Wörtwein, Louis-Philippe Morency, Ari Shapiro and Stefan Scherer. (2015) Exploring feedback strategies to improve public speaking: an interactive virtual audience framework: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp1143-1154.
- *Unity Game engine, tools and multiplatform.* (2017). Retrieved May 21, 2016, from https://unity3d.com/jp/unity
- Leap Motion. (2017). Retrieved May 21, 2016, from <u>https://www.leapmotion.com</u> About HATSUNE MIKU / CRYPTON FUTURE MEDIA. (n.d.). Retrieved May 21, 2016, from http://www. crypton.co.jp/mp/pages/prod/vocaloid/cv01.jsp
- *Tda formula Hatsune Miku Append Ver1.10 BowlRoll.* (2017). Retrieved May 21, 2016, from <u>https://bowlroll.net/file/4576</u> *piapro.net.* (n.d.). Retrieved May 21, 2016, from http://piapro.net/en_for_creators.html

Stereoarts Homepage. (2017). Retrieved May 21, 2016, from http://stereoarts.jp/ (in Japanese).