

A Tool to Transform Monologue Lecture Style into Dialogue Style by Adding a Tutee Agent in an Educational Video

Ari Nugraha^{†1} Izhar Almizan Wahono^{†1} Tomoyuki Harada^{†1} Tomoo Inoue^{†2}

Abstract : One interesting finding regarding lecture videos is that watching dialogue-style lecture videos produce better learning results for observer students than the monologue ones. Previous studies found dialogue-style lecture videos are better than monologue lecture videos because the observer student learns more from tutee than a tutor. Inspired by how the tutee plays a role in a dialogue video, we developed a system that transforms a monologue-style lecture video into dialogue-style lecture video by adding a tutee agent. As the result of the experiment of the system, most observer students preferred the dialogue style videos for various reasons.

1. Introduction

The majority of lecture videos provided in an online environment such as MOOCs, Open Courseware, and video-sharing platforms, are delivered in monologue style where one lecture/tutor presented the learning material [1]. Previous studies [1,2] found that dialogue observers student performing better than those given a monologue video, one of the reasons is the observers or students can learn as the tutee in the dialogue videos. Motivated by the tutee's role in dialogue video, in this study we proposed a system to enhance monologue style lecture video into dialogue-like style video to increase observer students learning performance and their attention spans when watching the lecture video.

2. Related Works

2.1 Video as a Learning Resource

Videos are still the primary resource of learning material in distance electronic education, particularly in MOOCs platforms. Popular MOOC platform such as Coursera, edX, Udacity, and Khan Academy made lecture videos as their main central course activity [3].

Student interaction with video material in MOOC platform have been explored by several studies. A large-scale study on video interaction patterns found that a high number of video replay and pause indicates a significantly higher perceived video difficulty [4]. Other study explored optimization of video interface such as providing preview or thumbnail of video parts where many learners tend to be interacted most [5].

Lecture video duration is also one of the topics explored by researchers in educational technology. One of the examples of this is empirical study conducted by Guo, Kim, and Rubin [6]

based on MOOC's engagements dataset. They found that shorter videos duration from few seconds to 6 minutes have higher engagement than videos with duration more than 6 minutes. Long lecture video duration tends to increase cognitive load of students, to reduce this, one of the strategies is to break the contents into smaller chunks of segments or segmentation. Breaking the learning content into smaller chunks of unit is called micro-learning, and it refers to learning processes cover a timespan from a few seconds to 15 minutes or more [6].

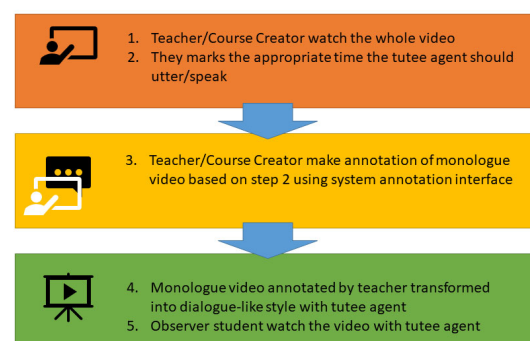


Figure 1. Workflow of Monologue to Dialogue Video Transformation

2.2 Learning Video Enhancement

As we mentioned before, segmentation can be used to break the lecture content into several segments. Short videos proven to have positive effect on learning outcomes, as study shown that learners in segmented groups have significantly better learning outcomes than in non-segmented group [7].

Recent study confirmed more on the effect of segmentation in procedural learning, where learners in the segmented interactive video group were better at performing the procedure than learners in noninteractive video groups [8]. Learning segmentation is not always producing positive effect to the learning process, like a too many segments are found to be

^{†1} Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan

^{†2} Faculty of Library, Information and Media Studies, University of Tsukuba, Japan

annoying and inappropriate [9]. Still in regard to segmentation, there are also exists studies related to automatically segments the lecture video, such as by analyzing the content of the synchronized slides using OCR [10] and analyze the lecture slides/presentation for any subtopic border such as title slide.

In-video quizzes, a small test which appears in the middle of video helps learners recall recently learned lessons [11]. A study found that in-video quizzes were successful in creating an engaging and interactive mode of content delivery [12].

Making annotation at a specific time point of the video where important concepts occur [13] also have been studied, as it is places minimal demands on the working memory and reduces disruption to the lecture as the student doesn't have to watch the entire video again [14]. Another type of marking specific scenes on video similar to annotation in the video have been explored. Researchers applied keyword tagging concept combined by gamification [15] to increase engagement. The that study, the participants tagging specific scenes on the video with keyword and got rewarded for their contribution.

2.3 Observational Learning

Observational learning, or observing learning activities carried out by others, such as observing dialogue between tutor and tutee proven to be better than learning by only observing tutor alone [1, 2]. An observer student, student who watch the video interaction between tutor and tutee, can master how to perform the activities without overt practice or direct incentives [16].

Several reasons found why dialogue-video better are, such as such as dialogue-observers pay more attention to what the tutees said than to what the tutors said, and tutees can serve as a model of learning [1].

Observational learning not only effective with human to human tutoring, it is also found that overhearing dialogue between virtual tutor and virtual tutee in virtual tutoring sessions have positives effects to observer student, such as student asked significantly more deep-level reasoning questions [17], and they can produce more relevant propositions in free recall questions [18].

3. Proposal

3.1 System Design

Inspired by previous studies on observational learning, learning segmentation and virtual tutoring sessions, in this study we want explore how virtual tutee agent could improve existing monologue video. Figure 1 visualized our system design workflow.

3.2 Implementation

To implement our proposed design, we developed our system which enables us to embed an avatar of tutee agent inside video

in web-based environment.

The tutee is embedded into the video dynamically using HTML5 Video API and JavaScript, in the form of an animated image and placed on the layer above the lecture video.

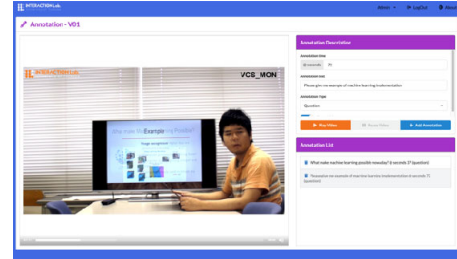


Figure 2. Tutee agent annotation interface

As seen in Figure 1, teacher or course creator needs to annotate their monologue video first using our annotation interface. The annotated texts created from the annotation interface by teacher are automatically converted into tutee agent dialogue, played at the designated time when the video is played. While the tutee agent is speaking, the other part of the video will be paused to mimic a dialogue between a tutee and a tutor.



Figure 3. Monologue video enhanced into dialogue-style with a tutee agent.

3.3 Hypothesis

Our hypothesis in this study is the observer students watching our tutee agent added lecture video will produce better learning performance than the observer students watching the monologue lecture video.

4. Initial Evaluation

4.1 Method

For the evaluation, we conducted a within-subject experiment involving 12 participants as an observer student with no prior experiences in using MOOC and unfamiliarity with the theme in the videos. Two conditions, Monologue (MON) and Dialogue with Tutee Agent (DIA) was used in this experiment. To counterbalance the effect of such as fatigue, learning effect, or bias from the learning style, Latin square design is used to make the order of the condition and the lecture video theme as shown

on Table 1.

Table 1. Experiment Conditions

	Lecture video themes	
	ID (Length)	MD (Length)
MON Condition	MON_ID (14:44)	MON_MD (13:28)
DIA Condition	DIA_ID (16:45)	DIA_MD (15:50)

As for the lecture videos, we produced two lecture videos in two themes, Interaction Design (ID) and Metadata (MD). The length of videos is similar, 14:44 minutes for ID theme and 13:28 for MD theme. All the dialogue of the tutor and the tutee agent inside the video are scripted. We used the talking-head style with slide for the video lecture-style, as it is proven to have better learning performance than other lecture-style [20]. We also used a different presentation slide style for each theme, where the ID theme contains more texts and explanations in each slide than the MD theme. And as for the dialogue style video, we used our system annotation interface to produce the videos by annotating the tutee agent to the monologue video. We designed the tutee agent role in the video as a summarizer and repeater of what the tutor said, where the tutee agent will do it at the end of each slide in the video like the example in Table 2.

Table 2. Example of tutee dialogue at the end of the slide in ID theme.

Tutor	"... For this second goal we need to take into account: Who the users are; What activities are being carried out; Where interaction is taking place..."
	At the end of current slide, before presentation moving to the next slide
Tutee	"So first we need to know the users are, second know their activities, and know where the interaction is taking place at"

4.2 Measurements

We measured several variables for this evaluation: learning gain by the number of relevant propositions from observer student's free recall/post-test quiz answer sheet (6 deep question for each theme), and observer's subjective rating on learning experience by questionnaire based on UES-SF [21] with 5 points Likert's scale ranging from strongly disagree to strongly agree for the answer.

4.3 Results

4.3.1 Learning Gain

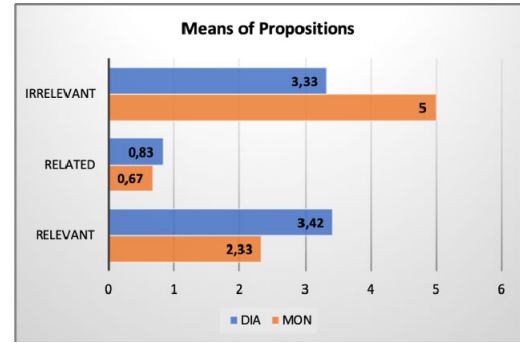


Figure 4. Means of propositions (N=12)

Figure 4 shows that DIA condition could produce better number of relevant propositions and smaller number of irrelevant propositions than MON condition. Learning result for each theme was also consistent with the overall learning gain result in Figure 4., where DIA produced better relevant propositions than MON in each theme.

Table 3. Means of relevant propositions for each video theme, MD (N=6) and ID (N=6)

	ID		MD	
	M	SD	M	SD
DIA	3.17	2.64	3.67	1.97
MON	2.67	2.66	2	2.45

4.3.2 Subjective Rating on Experience

As we have mentioned before, we used questionnaire after participants finished watching each video and answering the free recall test for the subjective rating on the participant experience when watching the lecture video. The results presented in Figure 5. showed that DIA could produce better user experience than MON, in all aspects except in the video difficulty, where DIA and MON gained the same rating with 2.5 out 5. These findings indicate that dialogue style (DIA) videos are more enjoyable, boosted confidence about the material, promoted focus attention, and more attractive than monologue style (MON) videos.

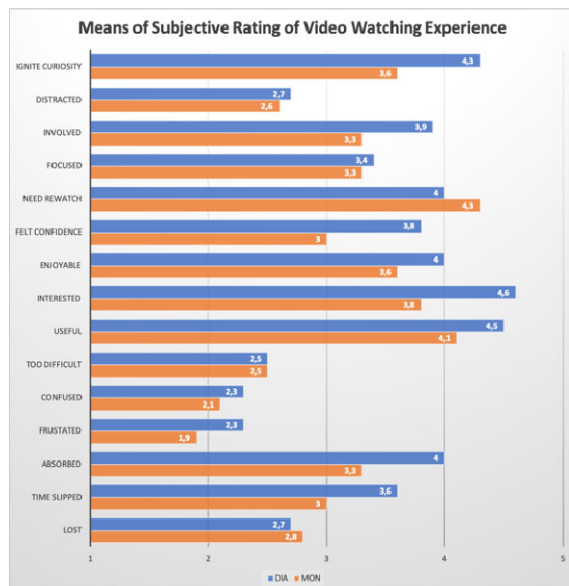


Figure 5. Means of subjective rating items

4.3.3 Qualitative Data

We gathered an interesting insight about what the participant's lecture style preferences and the reasons behind it from interview data. Eight (8) participants prefer dialogue style video, where three (3) participants prefer monologue style video, and one (1) participant does not have any preference for the lecture video style. The reasons for those participants who prefer dialogue style video are mainly because of the role of tutee agent as a summarizer of tutor's explanation like what P2 and P11 said in the interview, which in line with what intended for in this study.

"I found it easier for me to understand the video since the, there is an interactive dialogue in it, and also the lady (tutee agent) here, she said the term or the materials in the simple ways. Which I think, I found it useful for me..." [P2].

"The animated agent help to summarize what the lecturer says, so I can memorize better about the key point in the lecture, each of the subtopics, or sublecture (subtopic) of the whole lecture..." [P11].

5. Conclusion

Inspired by the benefits offered by dialogue style learning video, we propose a system which turns monologue video into a dialogue-like video by adding tutee agent to the monologue video. For evaluation we conducted experiments involving observer students to watch our dialogue-like and monologue video. Based on the data collected from free recall tests, questionnaires, and interviews with participants, we found that

our dialogue-like lecture video with the tutee agent could produce better learning and user experience than the monologue video.

References

- [1] Chi, M. T. H., Kang, S. & Yaghmourian, D. L. Why Students Learn More From Dialogue- Than Monologue-Videos: Analyses of Peer Interactions. *J. Learn. Sci.* 26, 10–50 (2017).
- [2] Muldner, K., Lam, R. & Chi, M. T. H. Comparing learning from observing and from human tutoring. *J. Educ. Psychol.* 106, 69–85 (2014).
- [3] Li, N., Kidziński, L., Jermann, P. & Dillenbourg, P. MOOC video interaction patterns: What do they tell us? *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 9307, 197–210 (2015).
- [4] Brinton, C. G., Buccapatnam, S., Chiang, M. & Poor, H. V. Mining MOOC Clickstreams: Video-Watching Behavior vs. In-Video Quiz Performance. *IEEE Trans. Signal Process.* (2016). doi:10.1109/TSP.2016.2546228
- [5] Kim, J., Li, S.-W. (Daniel), Cai, C. J., Gajos, K. Z. & Miller, R. C. Leveraging Video Interaction Data and Content Analysis to Improve Video Learning. *CHI '14 Work. Learn. Innov. Scale* 1–6 (2014).
- [6] Guo, P. J., Kim, J. & Rubin, R. How video production affects student engagement. *Proc. first ACM Conf. Learn. @ scale Conf. - L@S '14* 41–50 (2014). doi:10.1145/2556325.2566239
- [7] *The Cambridge Handbook of Multimedia Learning*. (Cambridge University Press, 2014). doi:10.1017/CBO9781139547369
- [8] Biard, N., Cojean, S. & Jamet, E. Effects of segmentation and pacing on procedural learning by video. *Comput. Human Behav.* (2018). doi:10.1016/j.chb.2017.12.002
- [9] Doolittle, P. E., Bryant, L. H. & Chittum, J. R. Effects of degree of segmentation and learner disposition on multimedia learning. *Br. J. Educ. Technol.* 46, 1333–1343 (2015).
- [10] Che, X., Yang, H. & Meinel, C. Lecture video segmentation by automatically analyzing the synchronized slides. 345–348 (2013). doi:10.1145/2502081.2508115
- [11] Kovacs, G. Effects of In-Video Quizzes on MOOC Lecture Viewing. in *Proceedings of the Third (2016) ACM Conference on Learning @ Scale - L@S '16* 31–40 (2016). doi:10.1145/2876034.2876041
- [12] Cummins, S., Beresford, A. R. & Rice, A. Investigating Engagement with In-Video Quiz Questions in a Programming Course. *IEEE Trans. Learn. Technol.* 9, 57–66 (2016).
- [13] Thomas, A. O., Antonenko, P. D. & Davis, R. Understanding metacomprehension accuracy within video annotation systems. *Comput. Human Behav.* 58, 269–277 (2016).
- [14] Moos, D. C. Note-taking while learning hypermedia: Cognitive and motivational considerations. *Comput. Human Behav.* 25, 1120–1128 (2009).
- [15] Viana, P. & Pinto, J. P. A collaborative approach for semantic time-based video annotation using gamification. *Human-centric Comput. Inf. Sci.* 7, (2017).
- [16] Craig, S., Driscoll, D. & Gholson, B. Constructing Knowledge from Dialog in an Intelligent Tutoring System: Interactive Learning, Vicarious Learning, and Pedagogical Agents. *J. Educ. Multimed. Hypermedia* 13, 163 (2004).
- [17] Cohen, P. A., Kulik, J. A., Kulik, C. C. & Cohen, P. A. Educational Outcomes of Tutoring : A Meta-Analysis of Findings. *Am. Educ. Res. Assoc.* 19, 237–248 (1982).
- [18] Driscoll, D. M. *et al.* Vicarious Learning: Effects of Overhearing Dialog and Monologue-like Discourse in a Virtual Tutoring Session. *J. Educ. Comput. Res.* 29, 431–450 (2005).
- [19] Craig, S. D., Gholson, B., Ventura, M. & Graesser, A. C.

Overhearing Dialogues and Monologues in Virtual Tutoring Sessions: Effects on Questioning and Vicarious Learning. *Int. J. Artif. Intell. Educ.* **11**, 242–253 (2000).

- [20] Ilioudi, C., Giannakos, M. N. & Chorianopoulos, K. Investigating differences among the commonly used video lecture styles. *CEUR Workshop Proc.* **983**, 21–26 (2013).
- [21] O’Brien, H. L., Cairns, P. & Hall, M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *Int. J. Hum. Comput. Stud.* **112**, 28–39 (2018)

Appendix

Appendix A.1 File history

Version number	Update details
V1.0	2019-12-20 First draft