SUGISAKA MASANORI

Journal of Advances in Artificial Life Robotics
Vol. 2(4); March (2022), pp. 216-219
ON LINE ISSN 2435-8061; ISSN-L 2435-8061
https://alife-robotics.org/jallr.html

Research Article

How does a TTS or highlighting system work for learning?

Sota Kobayashi, Masayoshi Tabuse
Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, 1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan

ARTICLE INFO

Article History
Received 25 November 2021
Accepted 25 April 2022

Keywords
Psychology of learning
Cognitive science
Verbatim memory
Text-to-speak
Highlighting

ABSTRACT
There are individual differences in human cognitive function, and that is a widely known fact. A hypothesis we made is that; giving both visual and audio stimuli may make it easier for people to catch information. In this research, three indicators are set consisting of memory, understanding, and concentration for an experiment. The difference in learning effect due to the reading situation was measured. We concluded that a Text-to-Speech (TTS) and the highlighting system can help reading in some cases.

© 2022 *The Author*. Published by Sugisaka Masanori at ALife Robotics Corporation Ltd.
This is an open access article distributed under the CC BY-NC 4.0 license
(http://creativecommons.org/licenses/by-nc/4.0/).

1. Introduction

There are individual differences in human cognitive function, and that is a widely known fact. For example, some people are good at reading text, others are good at listening. We made a hypothesis that; giving both visual and audio stimuli may make it easier for people to catch information. We conducted an experiment to verify the hypothesis.

2. Development Environment

The development environment is below.

- Language: C# (DotNet Framework 4.5)
- Integrated Development Environment: Microsoft Visual Studio Community 2019 (Version 16.8.2)
- Text to speak (TTS): CeVIO Creative Studio 7 (Version 7.0.23.1) [3]
- Morphological Analysis Engine: MeCab (MeCab.DotNet 0.0.26) [4]
- Heart Rate Monitor: Polar H10

3. Method

To verify the hypothesis, we made a system, which can highlight sentences and read them by the Text-to-Speech (TTS). The process is as follows.

1. Extract a target sentence from the full text. The target sentence is determined by punctuation marks.
2. The system passes the target sentence to CeVIO CS 7 and receives the pronunciation data. Then the system stores the data into a list structure whose members are mora and length of reading time.
   Mora is a unit in phonology. In Japanese, each kana character corresponds to a mora. Length of reading time is the time needed to read out the mora (msec).
3. The system passes the target sentence to MeCab and receives the morphological data. Then the system stores data into a new list structure named Manuscript whose members are morpheme, lexical category, and pronunciation. A morpheme is the smallest unit of the meaningful lexical item in language.
4. The system passes each morpheme to CeVIO CS 7 to check if the pronunciation is the same as that from MeCab and store the Boolean. If False, the system checks the adjacent node for the later processes. If the

Corresponding author's E-mail: kobayashi@mei.kpu.ac.jp, tabuse@kpu.ac.jp URL: https://www.kpu.ac.jp/

adjacent one is False too, they will be combined into one node.

5. The system stores length of reading time into each structure of the Manuscript. The lengths of reading times are calculated by referring to moraDataList. If the structure of the Manuscript has a True member, the length of reading time can be directly calculated. If False, the system finds out the next True and sums up the times before that.

   After this process ends, each node of the Manuscript has the following members: morpheme, lexical category, pronunciation, match of pronunciation (Boolean), and length of reading time (msec).

6. To make the Manuscript more comfortable to read, the software combines some structures. First, it combines by referring to the lexical category; if the word order was as determined, combines the structures (e.g., a postpositional particle is next after a noun). Second, it combines by referring to the reading time; if the reading time is shorter than 1000 msec, combines the structure with the next one.

7. Executing highlight; the system passes the Target sentence to CeVIO CS 7 to let it start reading out and then start highlighting.

4. Experiment

To investigate the effect of the differences in reading ways, three learning efficiency indicators were set: memory, understanding, and concentration. We referred to Terao's study [1] on memory and understanding, and Takatsu's study [2] on concentration.

4.1 Subjects

In this experiment, the subjects consist of 9 college students (male: A~F, female: G~I).

4.2 *Problem*

An experiment was conducted by using the problem set of [1]. The problem set consists of text and two elements: verbatim memory problems and understanding problems. The text is excerpted from a novel, an essay, or a fairy tale. They consist of around 270 characters.

In verbatim memory problems, the system shows 4 sentences to a subject and asks whether the same sentence was included in the text or not for each sentence. In understanding problems, the system shows 4 sentences to a subject and asks whether the sentence means were included in the text or not for each sentence.

As [1] said so, the difficulties of the problem set have been adjusted to be the same level.

4.3 Experiment

3 ways to read were set:

- read to oneself
  (hereinafter called SIL because read silently)
- read assisted by TTS
  (hereinafter called OFF because the highlight is off)
- read assisted by TTS + highlight
  (hereinafter called ON because the highlight is on)

Subjects read the sentences in 3 ways in random order no duplication and answered the questions; verbatim memory and understanding. They had been equipped with a Heart Rate Monitor 5 minutes before started solving the problem.

5. Results

Table 1 shows the scores of verbatim memory and understanding. Each problem set has 4 verbatim memory and 4 understanding problems. One point is given per one correct answer, therefore max is 4 points. Table 2 shows both the average and standard deviation of the scores.

Table 1. Score and average for each subject

Subject		A	B	C	D	E	F	G	H	I
memory	SIL	3	1	1	1	2	3	3	3	3
	OFF	4	2	2	2	2	3	2	2	3
	ON	2	3	2	1	3	3	2	2	1
understand	SIL	0	3	2	3	4	3	3	3	3
	OFF	2	1	1	2	2	2	4	2	4
	ON	3	1	2	2	2	4	3	3	3
average		2.33	1.83	1.67	1.83	2.50	3.00	2.83	2.50	2.83

Table 2. Ave and stdev of verbatim memory and understanding.

verbatim memory			understanding		
	ave	stdev		ave	stdev
SIL	2.22	0.916	SIL	**2.67**	1.05
OFF	**2.44**	**0.685**	OFF	2.22	1.03
ON	2.11	0.737	ON	2.56	**0.831**

Let's see Table 2. OFF got the highest average and lowest stdev scores of verbatim memory. It implies that OFF facilitates memorizing and prevents the influence of subjects' cognitive function. This result implies OFF can be an effective way to memorize.

SIL got the highest average of understanding which implies that SIL facilitates understanding (incidentally, ON got a relatively high score). ON got the lowest stdev of understanding which implies that ON prevents the influence of subjects' cognitive function. This result implies ON can be an effective way to understand.

Table 3 shows the score differences between ways of reading for each subject. Six correlation coefficients of rows were calculated; both memory and understanding have 3 combinations therefore 6 combinations exist.

Table 3. Score differences between the ways of reading for each subject

subject		A	B	C	D	E	F	G	H	I
memory	SIL - OFF	-1	-1	-1	-1	0	0	1	1	0
	OFF - ON	2	-1	0	1	-1	0	0	0	2
	ON - SIL	-1	2	1	0	1	0	-1	-1	-2
understand	SIL - OFF	-2	2	1	1	2	1	-1	1	-1
	OFF - ON	-1	0	-1	0	0	-2	1	-1	1
	ON - SIL	3	-2	0	-1	-2	1	0	0	0

Figure 1 and 2 shows the scatter plot of OFF-ON vs ON-SIL (memory) and SIL-OFF vs ON-SIL (understand), whose correlation coefficient absolute values are the biggest in 6 correlations.

As can be seen from Figure 1 and Figure 2, both certainly have a correlation. Then, we consider what the correlations mean.

(1) What does the correlation coefficient r = -0.763 for *OFF-ON vs ON-SIL (memory)* means?

① high OFF-ON: the score is decreased by highlight
② low OFF-ON: the score is increased by highlight
③ high ON-SIL: not good at memorizing by reading oneself but memory score is increased by highlight
④ low ON-SIL: good at memorizing by reading oneself but the score is decreased by highlight

The correlation coefficient r = -0.763 implies that ① &④, and ② & ③ have high positive correlations.

① & ④: who is good at memorizing by reading oneself could be impeded memory by the highlight.

② & ③: who is not good at memorizing by reading oneself could be facilitated memory by the highlight.

According to the above discussion, those who are good at memorizing by reading to themselves should not use highlights. On the other hand, highlights can be helpful to memorize for those who are not good at reading to themselves.

(2) What does the correlation coefficient r = -0.774 for *SIL-OFF vs ON-SIL (understand)* mean?

① high SIL-OFF: better at understanding by reading to oneself than reading with the system
② low SIL-OFF: better at understanding by reading with the system than reading to oneself
③ high ON-SIL: not good at understanding by reading to oneself but enhanced by highlight
④ low ON-SIL: good at reading to oneself but disturbed by highlight

The correlation coefficient r = -0.774 implies that ① & ④ and ② & ③ have high positive correlations.

For ① & ④: who are good at understanding by reading themselves could be not good at using the TTS and be disturbed by the highlight.

For ② & ③: who are not good at understanding by reading themselves could be good at using the TTS and be enhanced by the highlight.

According to the above discussion, those who are good at understanding by reading to themselves should not use highlight or TTS. Meanwhile, the TTS and the highlight could be useful for those not good at reading.

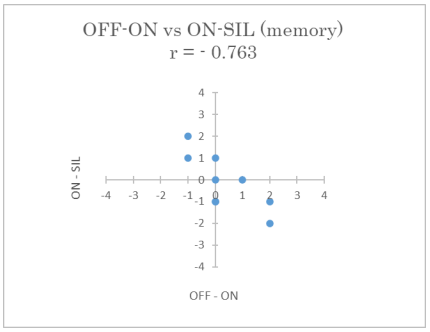

Figure 1. Scatter plot of OFF-ON vs ON-SIL (memory)

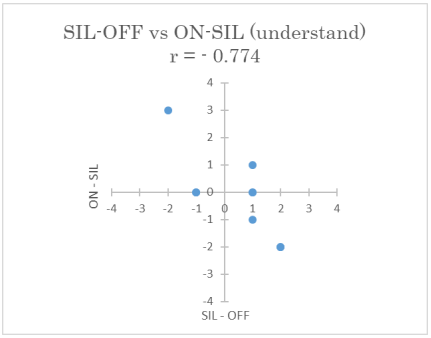

Figure 2. Scatter plot of SIL-OFF vs ON-SIL (understand)

Table 4 shows the personal average heart rate for each situation. Rest 1 is the lowest one-minute average while resting before solving the problem. Rest 2 is as same that but after solving the problem. REST in the bottom row is the smaller value of rest 1 and rest 2, which is the representative value. Whose reason is that some subjects can be nervous before or after the experiment. While nervous, the subjects are not at rest, so we took the lower one as a representative heart rate value of rest.

Table 4. Heart Rate (HR) for each subject (bpm)

subject	A	B	C	D	E	F	G	H	I
rest 1	86.8	97.9	80	101.9	71.3	62.3	87.4	78	71.9
SIL	86.5	106.1	81.1	103.5	77.2	70.1	99.3	81.4	72.3
OFF	90.3	97.9	82.5	102.3	79.6	69.3	97.1	83.5	72.4
ON	91.1	105.1	81.2	102.1	79.8	68.3	92.2	84.4	75.6
rest 2	87.6	97.6	81.7	97.3	78.2	67.2	94	80.5	70.1
REST	86.8	97.6	80	97.3	71.3	62.3	87.4	78	70.1

Table 5 shows the average and the standard deviation of each subject. Table 6 is the average and the standard deviation of the difference between reading and rest. As can be seen from Table 6. ON-REST has the highest average and lowest standard deviation. It implies that ON could be an effective way for those who don't have much concentration power.

Table 5. Ave and stdev of heart rate (bpm)

heart rate	ave	stdev
rest 1	81.9	12.1
rest 2	83.8	10.5
REST	81.2	11.5
SIL	86.4	12.7
OFF	86.1	10.9
ON	86.6	11.4

Table 6. Ave and stdev of difference between reading and rest

heart rate	ave	stdev
SIL - REST	5.19	3.70
OFF - REST	4.89	2.90
ON - REST	5.43	1.98

6. Conclusion

We researched the effects of the system that helps reading by using TTS and highlight. There are some cases where memorizing, understanding, or concentration are increased. This result implies that properly using TTS or highlight may help to learn.

7. Acknowledgement

I wish to acknowledge professor Yasunari Yoshitomi and specially appointed professor Taro Asada for useful discussions.

8. References

[1] T. Terao, M. Takahashi, S. Kiyokawa "Roles of articulatory movements and speech feedback in Japanese text comprehension during oral reading", The Japanese journal of psychology 89(6), 618-624, 2018. Shinrigaku-kenkyu. (In Japanese)

[2] H. Takatsu, O. Osamu "Assessment of Concentration Using Heart Rate Variability during taking class in college students", Journal of National Institute of Technology, Toyota College 39(0), 149-152, 2006 (In Japanese)

[3] CeVIO Project. Accessed 28 Jan. 2021

[4] MeCab.DotNet. Accessed 28 Jan. 2021

Authors Introduction

Mr. Sota Kobayashi

He received his B.S. degree from Informatics and Environmental Science, Kyoto Prefectural University, Japan in 2020. He is currently a master's course student at Kyoto Prefectural University.

Dr. Masayoshi Tabuse

He received his M.S. and Ph.D. degrees from Kobe University in 1985 and 1988 respectively. From June 1992 to March 2003, he had worked at Miyazaki University. Since April 2003, he has been at Kyoto Prefectural University. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are machine learning, computer vision, and natural language processing. IPSJ and IEICE members.