

Evaluation of a system that the reading of sentences by a voice synthesizer and the highlighting sentences

Sota Kobayashi, Masayoshi Tabuse

Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,

1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan

E-mail: s-kobayashi@mei.kpu.ac.jp, tabuse@kpu.ac.jp

<https://www.kpu.ac.jp/>

Abstract

It is generally known that there are individual differences in human cognitive function. We made a hypothesis; giving both visual and audio stimuli may make it easier for people to catch information. In this study, we set three indicators consists of memory, understanding and concentration for an experiment. The difference in learning effect due to the reading situation was measured. We concluded voice synthesizer and highlighting system can help reading in some cases.

Keywords: psychology of leaning, cognitive science, verbatim memory, voice synthesizer

1. Introduction

It is generally known that there are individual differences in human cognitive function. For example, some people have an advantage over reading the text, while others have an advantage over hearing the words. We made a hypothesis; giving both visual and audio stimuli may make it easier for people to catch information. To test this hypothesis, we conducted an experiment.

2. Development Environment

Development environment is as follows.

- Language: C# (DotNet Framework 4.5)
- Integrated Development Environment: Microsoft Visual Studio Community 2019 (Version 16.8.2)
- Voice Synthesizer: 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 test the hypothesis, we made a system that can highlight sentences and read by the voice synthesis. The process is as below.

- 1.Extract a target sentence based on the symbols like commas and periods that separate sentences.
- 2.The system gives the target sentence to CeVIO CS 7 to take pronunciation data. The system stores data into list structures whose members are mora and time-to-read. Mora is a unit in phonology. In Japanese, each character of kana corresponds to a mora. Time-to-read is the time to read a mora (msec).
- 3.The system gives the target sentence to MeCab to take Morphological data. The system stores data into new list structures named Manuscript whose members are morpheme, lexical category and pronunciation. Morpheme is the smallest unit of meaningful lexical item in language.
- 4.The system gives each morpheme to CeVIO CS 7 to check if the pronunciation is the same as that from MeCab. True or False will be followed by False join the next node other for the later processes.

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

After this process ends, each node of Manuscript has the following members: morpheme, lexical category, pronunciation, coordination-of-reading (boolean) and time-to-read (msec).

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

7. Executing highlight; give the Target Sentence to CeVIO CS 7 to let it start reading out and then start highlighting.

4. Experiment

To search for the effect of the differences in reading, we set 3 items of learning efficiency: memory, understanding and concentration. We referred to [1] on memory and understanding, and [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

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

In verbatim memory problems, the system shows a sentence to a subject and asks whether the same sentence was included in the problems or not. In the understanding problems the system shows a sentence to a subject and asks whether the sentences said the same meaning as the problem or not.

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

4.3 Experiment

We set 3 ways to read:

- read to oneself (hereinafter called SIL)

- read assisted by voice synthesis (hereinafter called OFF)
- read assisted by voice synthesis + highlight (hereinafter called ON)

Subjects read the sentences with 3 ways and answered the questions; verbatim memory and understanding. Subjects had been equipped with a Heart Rate Monitor 5 minutes before start to solve the problem.

5. Results

Table 1 shows the result of verbatim memory and understanding. Each sentence has 4 verbatim memory and 4 understanding problems. One correct answer gives one point, so max 4 points. Table 2 shows both the average and standard deviation of verbatim memory and understanding 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

As can be seen from 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. It means that OFF can be an effective way to memorize.

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

Table 3 shows the score difference between the way of reading for each subject.

Table 3. Score difference between the way 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

Six correlation coefficients of rows were calculated; both memory and understand have 3 combinations then 6 combinations exist.

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

As can be seen from the Figure 1 and 2, each of them certainly has a correlation. Here, we consider what the correlations mean.

(1) Correlation coefficient $r = -0.763$ for OFF-ON vs ON-SIL (memory)

- ① Subject has high OFF-ON: disturbed by highlight
- ② Subject has low OFF-ON: enhanced by highlight
- ③ Subject has high ON-SIL: not so good at read to oneself but enhanced by highlight
- ④ Subject has low ON-SIL: good at read to oneself but disturbed by highlight

The negative correlation coefficient implies that ①,④ and ②,③ have high positive correlations.

For ①,④: who good at read to oneself could be disturbed by highlight.

For ②,③: who not good at read to oneself could be enhanced by highlight.

According to the above discuss, who good at memorize by rearing to oneself should not use highlight. On the other hand, highlight can be helpful to memorize for who not good at read to oneself.

(2) Correlation coefficient $r = -0.774$ for SIL-OFF vs ON-SIL (understand)

- ① Subject has high SIL-OFF: better at reading to oneself than reading by using reading system

- ② Subject has low SIL-OFF: better at reading by using reading system than reading to oneself
- ③ Subject has high ON-SIL: not so good at read to oneself but enhanced by highlight
- ④ Subject has low ON-SIL: good at read to oneself but disturbed by highlight

The negative correlation coefficient implies that ①,④ and ②,③ have high positive correlations.

For ①,④: who good at read to oneself could be not good at using reading system and be disturbed by highlight.

For ②,③: who not good at read to oneself could be good at using reading system and be enhanced by highlight.

According to the above discuss, who good at understand by reading to oneself should not use highlight or reading system. On the other hand, reading system and highlight could be useful for who not good at reading.

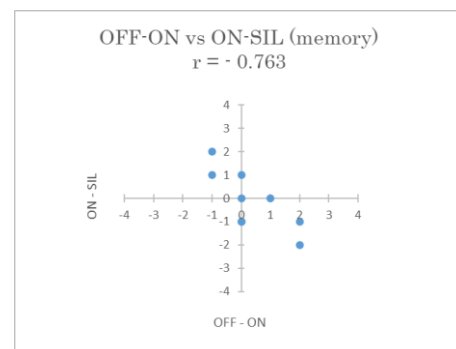


Figure 1. Scatter plot of OFF-ON vs ON-SIL (memory)

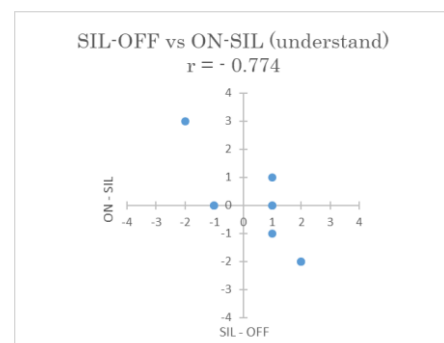


Figure 2. Scatter plot of SIL-OFF vs ON-SIL (understand)

Table 4 shows the personal average heart rate for each behavior. Rest 1 is the lowest one minute's average while being rest before solving the problem. Rest 2 is as same to that but after solved 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 subject can be nervous before or after experiment. While in nervous, the subjects are not in rest, so we took lower one as 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 standard deviation of each subject. Table 6 is the average and standard deviation of difference between reading and rest. As can be seen from Table 6. ON-REST has highest average and lowest standard deviation. It implies that ON could be an effective way for who doesn'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 investigated the effects of the system that helps reading by a voice synthesis and highlighting. There are cases that memorizing, understanding or concentration are increased. This result implies that properly using voice synthesis or highlight may help learners.

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, <https://cevio.jp/>, Accessed 28 Jan. 2021
4. MeCab.DotNet, <https://github.com/kekyo/MeCab.DotNet/blob/master/R/EADME.ja.md>, 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 course student of 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 in Miyazaki University. Since April 2003, he has been in 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 member.
