International Journal of Advances in Computer Science & Its Applications – IJCSIA

Volume 9 : Issue 2- [ISSN : 2250-3765] - Publication Date: 27 Dec, 2019

# System Design of Multi Lingual Mobile Rapid Sentence Reading

Toshiyuki Maeda

***Abstract*— This paper presents a system design of multi-lingual enhanced Rapid Serial Visual Presentation (RSVP) system "MulLinR," which supports languages written continuously (non-spaced) and spelled with nonalphabetic characters. Our prototype accepts Chinese and Japanese for the time being. We express the system design of our prototype. One of the features of the system is that it is fully written in JavaScript, and so it is a standalone application without a network connection. For the separation of words, we introduce "TinySegmenter" which use feature quantity of characters and their combinations as N-gram, character kinds such as Kanji, Kana, etc. We discuss the experiment with a variety of presentation time intervals depending on the word length.**

***Keywords*— Rapid Serial Visual Presentation, Multi Lingual, presentation time intervals, word length**

## I. Introduction

Mobile phones are one of the most important gadgets for mobile communication in these days. Using small displays of mobile phones, we can hardly read in a usual way of "from left to right" and "from top to bottom" orders. "Rapid Serial Visual Presentation" (RSVP) [7] can be a useful method to read long sentences.

"Spritz" is one of the most well-known application of RSVP, which displays only one word at one time respectively from the begging to the end of sentences [9]. Spritz presents "focused reading experience and help readers get their content faster, with less effort" [9].

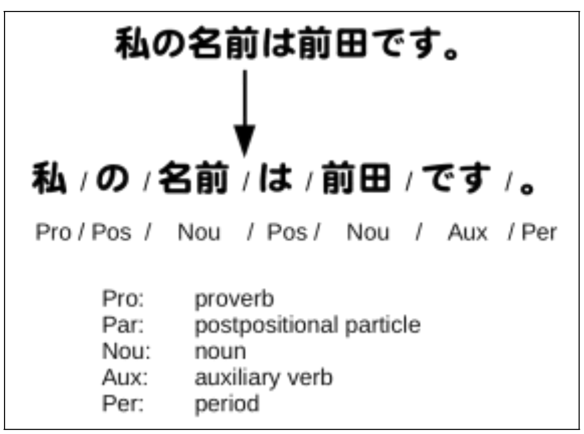


Figure 1: An example of test separation

Toshiyuki Maeda
Hannan University
Japan

However, one of the most important problems for Spritz is that Spiritz is almost focused on English or other languages spelled by Roman alphabets. Non-alphabetic characters languages such as Chinese, Japanese, and so on, are not supported. There is, moreover, another important problem in some languages. For example, in some languages, words are not separated by spaces but written continuously, and so the system has to disconnect words into one or a few words before displaying like RSVP as in Figure 1.

Some researchers try to solve the problems in Japanese, for instance, [2]. Nevertheless, the system depends on Yahoo Web API [11] and then it must have some latency. That requires, furthermore, connectivity of Internet.

We developed an RSVP system called "MulLinR", which resolves some of above language-specific problems, and discuss the usability [6].

In this paper, we explain the system design of our improved system, which can configure the interval in cooperation with the length of a word, and then we carried out experiments and discuss the results.

## II. Related Works

Many researchers investigate RSVP so far. Reeves et al. [8] claim that Three graduate students with normal or corrected visions were investigated the shift of attention to the sentences by rapid sentence visual representation. The sentence detected the target embedded in the series of characters displayed on the left side of the fixation, and as soon as possible shifted attention to the series on the right side of the fixation. They attempted to report, in order, the 4 earliest occurring numerals after the target. They analyze for all sentences, targets, and many rates, the relative position of numerals in the response sequence showed clustering, disorder, and folding. The numbers reported indicated to cluster around the stimulation location 400 ms after the target. The numbers were clearly reported in the order of chance. The actual reporting order is a mixture of correctly displayed numbers and numbers in the opposite direction of the display order. The results are described quantitatively by order strength theory and are efficiently predicted by the computational attention gating model (AGM). The AGM may be derived from a more general attention model that assumes that after detection of the target, an attention gate opens briefly to allow numerals to enter a visual short-term memory, and subsequent order of report depends on both item strength (how wide the gate was open during the numeral's entry) and on order information (item strength times cumulative strength of prior numerals). On the Spritz site [10], traditional reading involves aligning text and moving the user's eyes in order in turn. For each word, the eye looks for a specific point in the word. This is called the "Optimal Recognition Point" or ORP. After their eyes find the ORP, their brains begin to process the meaning of the words they are looking at. Every time a new word called "Saccade" comes out, the eyes move, searching for the word ORP and moving there. As their eyes focus on the ORP, they begin to process words for meaning and context, and their eyes move on to the next word. When their eyes encounter punctuation in and between sentences, their brains are urged to collect all the words they read and process them into coherent thoughts. Thus, Spratzing offers to read the content one word at a time, with the ORP placed in our conventional "redicle." Some of the redicles have special prompts that allow them to read without having to move their eyes and keep their eyes focused on the ORP. Placing words exactly where your brain wants is an intensive, engaging and enjoyable reading experience.

## III. System Design

As we mentioned, the most important function for MulLinR is that sentences written continuously must be separated by word(s) in a certain length of characters. It must be, besides, preceded in a reasonable manner as shown below. We develop morphological analysis modules to arrange or modify words for improving readability, and after that the words are displayed as Figure 2.

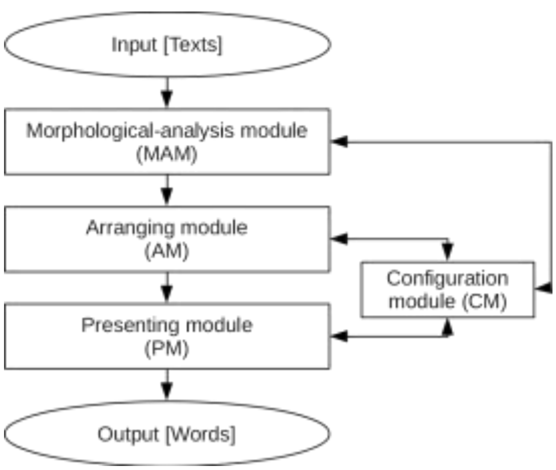


Figure 2: System overview.

We have developed a prototype system for MulLinR, and that supports Japanese and Chinese (See screen shots in Figure 3, 4). This prototype system uses "TinySegmenter" [5] for implementation of MAM as for Japanese. TinySegmenter does not use dictionaries because of tiny system design. TinySegmenter writes a Japanese sentence leaving some space between words or grammatical units, only using machine learning data. TinySegmenter learns and classifies all characters of inputed sentence whether a character is preceded by a word boundary or not, using feature quantity of characters and their combinations as N-gram, character kinds such as Kanji, Kana, etc. In some cases, analysis accuracy of unknown words is better than by MeCab as of not using dictionaries. This process is basically language-independent, and so we can make analysis models regardless of languages if we have many space-separated sentences.


Figure 4: MulLinR screen shot (for Japanese).


Figure 4: MulLinR screen shot (for Chinese).

As mentioned above, TinySegmenter is originally only for Japanese, and then we improve for Chinese implementation. We introduce "TinySegmenterMaker" [4] to construct analysis model of space-separated sentences in Chinese for TinySegmenter. TinySegmenterMaker uses AdaBoost algorithm to learn space-separated sentence data. AdaBoost, or Adaptive Boosting, is a meta-algorithm of machine learning. It is used in combination of some other learning algorithms to improve performance. In order to make learning data of Chinese sentences, "friso" [3] is introduced as morphological analysis in Chinese.

## IV. Experiment and Discussion

In [6], we carried out the experiment and got the result that 400 milliseconds interval is the best in our experiment. In this experiment, the improved supports to modify the interval in cooperation with the length of a word such as;

$$I = a \times L + b$$

where I is the variation of interval (ms) for the word, a is the parameter of word-length dependent constant, L is the length of the word, and b is the parameter of base interval length constant. It is because we assume the above equation as, for instance in Japanese, Kanji characters have their own meanings and so require some sort of time so as to be read and understood. We have experiments so as to discover fitting time parameters for MulLinR only for Japanese. As for Japanese readings, 11 to 25 subjects (depending upon cases) read the sentences, Chapter one of "Kumo no ito (Spider's Thread) " written by Ryunosuke Akutagawa [1].

| [*a*/*b*] | [200/50] | [200/100] | [100/100] | [400/100] |
|---|---|---|---|---|
| Very good | 1 | 7 | 6 | 7 |
| Good | 14 | 4 | 5 | 2 |
| Too fast | 6 | 0 | 8 | 4 |
| Too late | 3 | 0 | 6 | 12 |

Table I: Experiment result.

First of all, although it is possible to read the character with the default speed [200/50], it does not seem the best interval. We got a result that more than half could read normally, and it seemed that it was a well-balanced value at first glance, but when they continue to read, then some of them may feel tired. Few of them evaluate "Very good," and the opinion of "fast" was most frequently seen by female subjects in this experiment, and so it is not suitable to recommend [200/50], especially to those who use for the first time. In terms of the next target [200/100], the number of character dependencies increased by 50 from the default, and it seems easier to read "Good." However, since the experiment was conducted after verifying the other three values for a while, the number of subjects decreased compared to the other values. Although the data is small in number of subjects, all the subjects could read more than usual, and it seems quick or slow for various subjects, and so we this value [200/100] may be recommended for the first time.

Moreover, personally satisfying values [100/100] resulted in the most divergent of the four tests. There were many subjects who felt fast while there were a certain number of subjects who felt slow, and we tried to find out about this gap of awareness by ourselves. That is the odd feeling of the character dependency. First, by setting the base speed to be 100 less than the default comparison target [200/100], the speed at which the characters are switched is originally faster, and a subject who feels faster than the default comparison target results However, it should be strange that some subjects may feel late. However, the fact that the subject felt slow was that something felt slow, that is, because of the speed of the fast-flowing character, the time during which the character dependency stopped was felt long and felt late It is. Because the base speed is half of the default comparison, the differences in the letter dependencies are visible and noticeable. In the case of me, it takes a long time to read long sentences because we usually do not read books very much, but it is a type that feels easier to read if there are many character dependencies. Readability may be affected by not being fit for me. It is not recommended for people who are accustomed to reading sentences for the first time as to whether it can be recommended for those who touch the text speed reading application for the first time, but for people who take longer to read one letter like me we felt it was a recommended value.

The value that the first set of men felt easy to read [400/100] felt that many subjects felt relatively "slow." We also felt late about this, but there were four men who felt fast. We thought that all subjects who felt just right with the previous values should feel late because the base speed is more than double that of the other values, but the four people here also have a sense of incongruity due to the character dependency mentioned earlier we thought that we felt fast. However, in addition to the first set of men, there were many subjects who felt that this speed was just right, so it is a highly recommended value for people who are slow to read characters or who want to read characters carefully.

We experienced to some extent a letter display speed reading application and thought about the tips for tracking letters smoothly when using it. To that end, we felt that it was important to read the displayed letters without speaking out in my head. If you read letters one by one in your head when you read letters you saw with your eyes, it will take time to be recognized by your brain, so it is better to recognize the letters one by one or to read them quickly. We are able to read the sentences smoothly and quickly while we are reading.

# V. Conclusion

We address a multi-lingual RSVP system and develop a prototype in order to clear up languagespecific problems. We explain the system design and the implementation of MulLinR and have experiments and discuss the time intervals variation.

We need more experiments as those experiments had not got enough scale of result data, for example, smaller time interval variations, and so forth.

## *Acknowledgment*

Part of this research was supported by JSPS KAKENHI Grant Number 18K02920 and 19K03018. This research was also partially assisted, especially for carrying out the experiment, by Mr. Hidehiro Nakanishi at Hannan University. The authors greatly appreciate the supports.

## *References*

[1] R. Akutagawa. Kumo no ito (in Japanese). https://www.aozora.gr.jp/cards/000879/files/92_14545.html, 1918 (Retrieved in April, 2019).

[2] Furyu-tei. Japanized pseudo-spritz (in japanese). http://furyu.hatenablog.com/entry/20140325/1395741983, (Retrieved in June, 2016).

[3] Lionsoul2014 (github). friso. https://github.com/lionsoul2014/friso, (Retrieved in July, 2019).

[4] Shogo82148 (github). Tinysegmentermaker. https://github.com/shogo82148/TinySegmenterMaker, (Retrieved in July, 2019).

[5] T. Kudo. Tiny segmenter (in japanese). http://chasen.org/~taku/software/TinySegmenter/, (Retrieved in June, 2016).

[6] T. Maeda, Y. Fukushige, M. Yajima, and A. Wakatani. Multi-Lingual Pseudo Natural Language Programming Environment for Beginners. In *Proceedings of World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2018)*, page (at www.academicexperts.org/conf/edmedia/2018/), Amsterdam (Netherland), 2018.

[7] M. C. Potter. Short-term conceptual memory for pictures. *Human Learning & Memory*, 2(5):509–522, 1976.

[8] A. Reeves and G. Sperling. Attention gating in shortterm visual memory. *Psychological Review*, 93(2):180– 206, 1986.

[9] Spritz. Spritz homepage. http://spritzinc.com/, (Retrieved in June, 2016).

[10] Spritz. Spritz the science. http://spritzinc.com/thescience, (Retrieved in June, 2016).

[11] YahooJapan. Yahoo web api (in japanese). http://jlp.yahooapis.jp/DAService/V1/parse, (Retrieved in June, 2016).