Non-native Language Reading Support with Display of Machine Translation Based on Eye-Tracking and Sentence-Level Mapping

Tien-Yu Ho
National Tsing Hua University
Hsinchu, Taiwan
skyand55@gmail.com

Hao-Chuan Wang
National Tsing Hua University
Hsinchu, Taiwan
haochuan@cs.nthu.edu.tw

Shong-Hong Lai
National Tsing Hua University
Hsinchu, Taiwan
lai@cs.nthu.edu.tw

ABSTRACT
In the era of information explosion, individuals often need to read first-hand information online in English, which is a non-native language of most of the people in the world. However, the lack of proficiency toward a non-native language makes it difficult for second-language readers to efficiently understand the contents on the English webpages, especially the long articles. In this work, we propose a novel reading interface to support English as second-language readers with adaptive display of machine translation (MT) using eye tracking, accompanied with sentence-level mapping using background colors. We conducted an experiment to investigate how different methods affect the second-language readers. We found that active translation could help second-language readers comprehend the English article without dazzling them. In addition, sentence-level mapping using background colors not only benefits the mapping between original sentences and their corresponding ones, but also alleviates the problems of line skipping and reading resumption.

Author Keywords
Non-native language; eye tracking; reading support; machine translation.

INTRODUCTION
As information interfaces and services become prevalent on Internet, information access becomes less of a problem to global users. However, information we want on the web is not always written in our first language. In fact, since, much of the information were available in English, which is a second language for most of the people in the world. Non-native English readers tend to encounter reading comprehension problem, thus limiting their work efficiency.

ChineseCHI '18, April 21–22, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-6508-6/18/04$15.00
https://doi.org/10.1145/3202667.3202675

In this paper, we intend to assist people whose second language is English to enhance their comprehension of English articles. As we're interested in supporting reading without relying on multimodal input, we choose to focus on eye-tracking in this paper to control the display of machine translation (MT) results. In addition, we propose *sentence-level mapping*, which separates a paragraph into sentences for better readability and translation mapping, and uses different color codings to alleviate the reading resumption problem caused by attention switching. A sentence in the original paragraph has the same background color as its corresponding translated sentence.

BACKGROUND
For years, researchers have tried to reduce the language barrier in reading with many different ways. Yu et al. [15] proposed a new paragraph transformation approach called Jenga Format (JF) to enhance user satisfaction and reading comprehension for non-native readers by inserting white spaces between sentences. Picture note-taking [3] was also proved to be a feasible way to help second-language readers to comprehend English paragraphs.

Eye tracking has been applied to research related to reading behavior and interaction. Okoso et al. [11] utilized the annotation of reading speed, re-reading and number of fixation areas to determine English learners' text comprehension level toward English articles. Hohlfeld et al. [4] presented a system that supports the diagnosis and therapy of reading disorders by mobile gaze tracking. Researchers have investigated how to improve reading comprehension by using the gaze tracking technique. Copeland et al. [1] investigated how the eye movement factors influence reading and overall comprehension. iDict [5] is a translation-assisted design for foreign language reading that detects user's eye movement to understand the demand for help. Gazemarks [8] and EyeBookmark [6] made use of reader's gaze point to provide a clue for reader to restore to the original task rapidly on web pages.

There are also some works devoted to the study of gaze control, especially triggering. "Target reverse Crossing" was proposed by Feng [2], as a gaze-only selection method, which was afterwards applied to EyeSwipe [9], a dwell-time-free gaze-typing method. Lutteroth et al. [10] presented another gaze-only click alternative called Actigaze, which marks the potential targets with different background colors and shows some corresponding colored confirmation buttons. Confirming to click a button can be simply done by color matching. They proved that the Actigaze can effectively reduce the clicking time compared to Multiple Confirm click alternative [12], which does not have color mapping.

EXPLORATORY STUDY

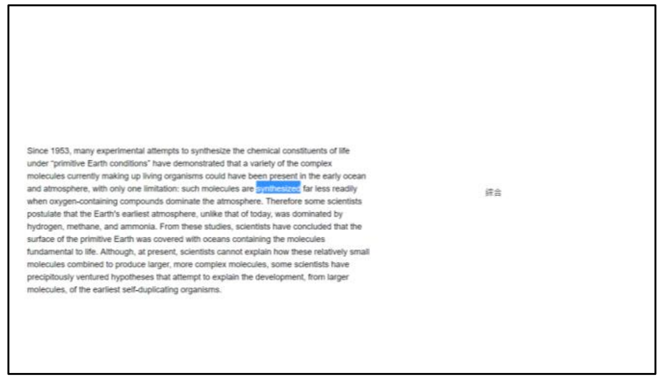
For the sake of figuring out what kind of native-language support that makes users easier to read and understand the second-language text, we conducted a within-subject study with 18 participants whose native language is Mandarin and second language is English.

This study consisted of three conditions, L1, L2, and L3, which denote paragraph-level, sentence-level, and word-level translation, respectively. In each condition, the participants were required to read an English paragraph chosen from GRE mock test with a specific first-language supporting method in 90 seconds, and answer a reading comprehension question with no time limit. Participants could not re-read the paragraph when answering the question. The paragraphs in the three conditions were different, and the order of the conditions and paragraphs were counterbalanced by Latin Square.

The layout for each condition composed of three parts: countdown timer in seconds on the left-top corner, the original paragraph on the left-hand side, and the translated content on the right-hand side, as depicted in Figure 1.

(a) L1: Paragraph-level translation

(b) L2: Sentence-level translation



(c) L3: Word-level translation

Figure 1. The user interfaces of the three conditions in exploratory study.

Results

We measured participants' correctness of the answer of comprehension test and preference toward the three different supporting methods. The results are shown in Figure 2 and Figure 3.

We found from the experimental results that *automatic sentence separation* and *sentence-level translation mapping* are important to second-language readers. However, we also found that the sentence separation method we used in the sentence-level translation condition of preliminary study interrupted reading continuity. As a result, we modified the sentence separation and mapping method by using different background colors for each sentence in the original text and its corresponding translated text.

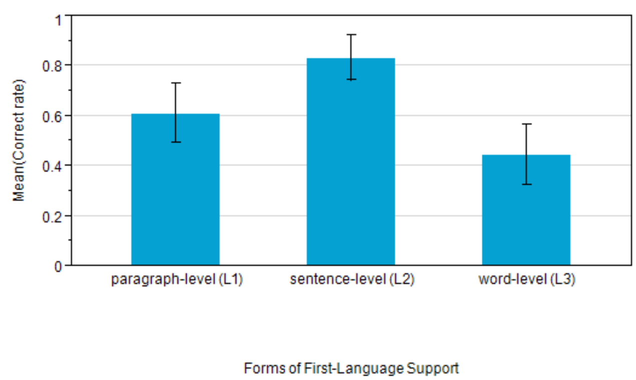


Figure 2. The correctness rate for different forms of first-language support.

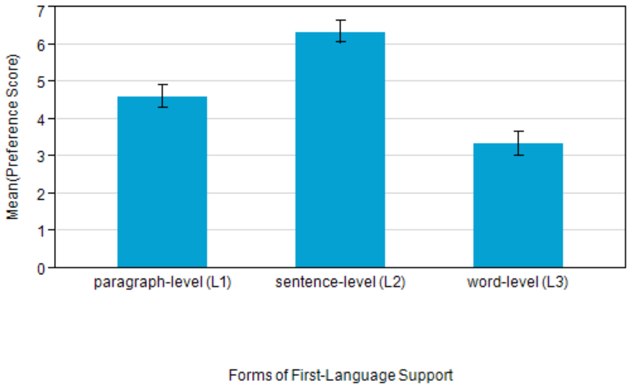


Figure 3. Participants' preference scores for different forms of first-language support.

MAIN STUDY

In this study, we would like to explore how different MT display methods and different sentence separation and mapping methods affect readers' comprehension and readability of second language articles. The goal was to allow second-language reader to easily map the corresponding translation of the original text when they read articles. The interface we introduced is active display of MT adopting eye tracking technique, as well as sentence separation and mapping method using background colors. It is the first condition (C1) in the experiment. Figure 4 shows the screenshot of the proposed interface.

Since there were too many conditions for a single participant, we divided all the participants equally into two groups, G1 and G2, by the principle of maintaining the balance of English ability between two groups. Therefore, each participant within a group would only need to accomplish the tasks in three conditions so that they would not be too tired when reading articles. We used webpages as the platform to conduct the experiment.

Figure 4. Screenshot of the proposed reading interface. The orange dot represents user's gaze position.

Participants

We recruited 18 participants for the study whose native language is Mandarin and second language is English. All of the participants took a language proficiency test from Cambridge English Language Assessment [14] first and reported their scores. We divided all the participants equally into two groups (G1 and G2) by the scores, so that the English abilities of two groups are similar ($M_{G1} = 18.44, SD_{G1} = 1.88$; $M_{G2} = 18.44$, $SD_{G2} = 2.29$).

Tasks

There were six conditions in total (C1 to C6), which was the combination of three display of MT levels and two sentence-level mapping levels. The three levels of display of MT were t1, t2, and t3, which were active translation (with eye tracker), passive translation and no translation, respectively. And each of the display level was supported with/without sentence-level mapping (s1/s2). The details of each condition are listed as follows: (C1) active translation with sentence-level mapping, (C2) passive translation with sentence-level mapping, (C3) no translation with sentence-level mapping, (C4) active translation without sentence-level mapping, (C5) passive translation without sentence-level mapping, (C6) no translation without sentence-level mapping.

Active translation	Using eye tracker for detection of user's gaze position. When the user gaze at a specific paragraph, only the corresponding content would be translated and displayed at the same vertical position. The cursor would be hided.
Passive translation	The entire article would be translated and displayed.
No translation	The entire article would not be translated.
Sentence-level mapping	Paragraph(s) would be automatically divided into sentences. Each translated (if any) and corresponding sentences pair would be marked with the same background color.
No sentence level mapping	There were no division nor coloring to all the contents.

Table 1. Introductions of the three display of MT levels and the two sentence-level mapping levels.

All the participants in G1 would be supported with sentence-level mapping (C1 to C3), while those in G2 would not (C4 to C6). Table 1 is the introductions of the three display of MT levels and the two sentence-level mapping levels. The webpage layout of each condition composed of three parts: a button for participants to click when they had done their reading on the left-top corner, the original article on the left-hand side, and the translated content on the right-hand side, as shown in Figure 5. The original articles and the following questions are chosen from previous test of TOEFL provided by [7]. For each article, we modified the original one to three paragraphs, and the translated texts were translated by Google Translate API. G1 and G2 shared the same three articles.

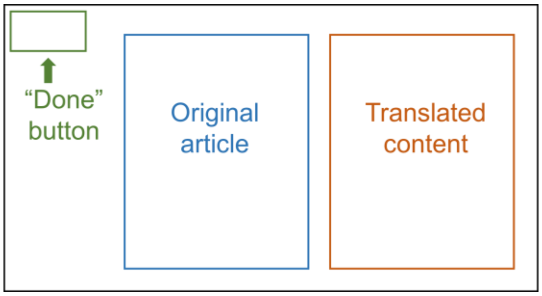


Figure 5. The diagram of the reading interface.

Software and Equipment

The experiment environment was set in a quiet room and would not be disturbed during the experiment. The setting consisted of three main parts: a laptop (Windows 10, 12GB RAM, i7 2.6GHz CPU), an external monitor (1920x1080P, 24”), and an eye tracker (Tobii EyeX, 60 Hz) attached to the bottom of the external monitor. The distance between the monitor and user is about 60cm. We also adopted Project IRIS [13], which allows users to control the mouse using gaze points detected by Tobii EyeX and hide the cursor. The webpages used for reading and displaying were implemented by JavaScript.

Procedure

Participants were assigned to G1 or G2 according to their scores of Cambridge English exam. At the beginning of the experiment, the experimenter introduced the study and explained the tasks and the reading interface to the participants. For G1, we explained the content of C1, C2, and C3. As for G2, we explained C4, C5, and C6 to them. Next, the calibration session would be executed to ensure the accuracy of the gaze points detected by eye tracker. Then the participants would have about 5 minutes to get familiar with the reading interface and the usage of eye tracker.

Afterwards, the test session was started. The test session was made up of three article readings with different combinations of three display levels of MT levels and two sentence-level mapping levels, each of which followed by three reading comprehension questions and one online survey about the reading experience. All the parts in the test session were done without time constraint. We also made an interview afterwards to understand their reading experience and their opinions toward the three conditions. The experiment took about 75 minutes per participant.

MEASURES

We used participants’ subjective opinions toward their reading experience and one objective measure to evaluate the two factors: display of machine translation and sentence-level mapping.

Correctness rate

We used the proportion of correct answers among three multiple choice questions of the comprehension test as an objective assessment. The accuracy rate was served as the objective index to the level of their reading comprehension.

Presentation readability

We measured the participants’ feeling about presentation readability of the interfaces with 7-point Likert scale: *(1) I sometimes skip lines while I was reading, (2) I was dazzled while I was reading.*

RESULTS

Display of MT and sentence-level mapping are the two independent variables in this experiment. We used two-way ANOVA to analyze the collected data.

Correctness rate

There was significant difference in display of MT: $F[2,48] = 3.9407, p < 0.05$. The post-hoc Tukey HSD Test showed that participants’ objective comprehension was significantly higher when using active translation ($M_{t1} = 0.6297, SD_{t1} = 0.3001$) and passive ($M_{t2} = 0.6481, SD_{t2} = 0.2418$) translation over no translation ($M_{t3} = 0.4259, SD_{t3} = 0.2506$): $p < 0.05$. Participants’ correctness rates were not significantly different between reading with ($M_{s1} = 0.6296, SD_{s1} = 0.2824$) and without ($M_{s2} = 0.6296, SD_{s2} = 0.5061$) sentence-level mapping: $F[1,48] = 2.7197, p = 0.0916$. There was also no significant interaction effect between display of MT and sentence-level mapping: $F[2,48] = 0.3852, p = 0.6824$. Figure 6 shows the means and standard errors of correctness rates for different levels in the display of MT and sentence-level mapping.

Presentation readability

Line skipping

For line skipping problem, we did not find any significant difference among active ($M_{t1} = 3.1667, SD_{t1} = 1.9174$), passive ($M_{t2} = 3.5556, SD_{t2} = 1.9770$) and no translation($M_{t3} = 3.0556, SD_{t3} = 1.6618$): $F[2,48] = 0.4288, p = 0.6538$. As we expected, participants felt significantly much more likely to skip lines unintentionally when they were reading without sentence level mapping ($M_{s2} = 4.0741, SD_{s2} = 1.9400$) than reading with sentence level mapping ($M_{s1} = 2.4444, SD_{s1} = 1.3107$): $F[1,48] = 12.3904, p < 0.001$. No significant interaction effect was found between display of MT and sentence-level mapping on the score of subjective agreement toward the statement “I sometimes skip lines while I was reading”: $F[2,48] = 0.1984, p = 0.8207$. Figure 7 shows the means and standard errors of the scores of line skipping for different levels in display of MT and sentence-level mapping, respectively.

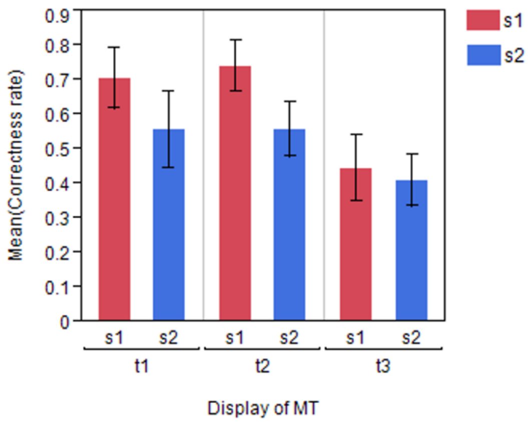


Figure 6. Correctness rates by conditions.

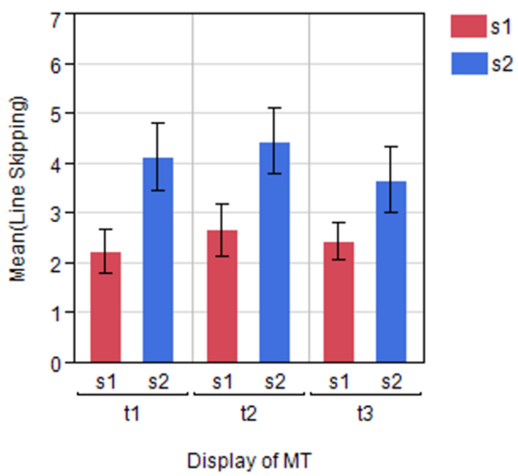


Figure 7. Line skipping by conditions.

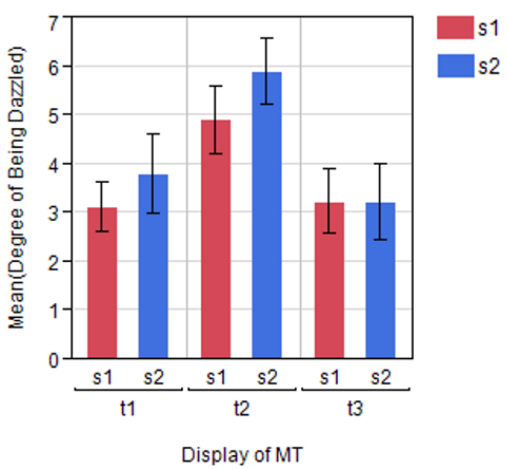


Figure 8. Degree of being dazzled.

Degree of being dazzled

The mean scores and their standard errors for different levels of display of MT as well as sentence-level mapping were shown in Figure 8. The main effect of display of MT was significant: $F[2,48] = 5.8641, p < 0.01$. The post-hoc Tukey HSD Test provided that the participants' degree of being dazzled was significantly higher by using passive translation ($M_{t2} = 5.3889, SD_{t2} = 2.0619$) over no translation ($M_{t3} = 3.2222, SD_{t3} = 2.1020$): $p < 0.01$, and passive translation over active translation ($M_{t1} = 3.4444, SD_{t1} = 2.0065$): $p < 0.05$. No significant difference was found for sentence-level mapping on this measure: $F[1,48] = 0.9554, p = 0.3332$. There was no significant interaction effect between display of MT and sentence-level mapping as well: $F[2,48] = 0.2675, p = 0.7664$.

DISCUSSION

The current eye-tracker we used in the prototype is a consumer-grade model, which may limit the performance of active translation support. However, we still see signals of usefulness of this support, showing the potential value of this design. Improvement in eye-tracking technologies can possibly lead to enhanced outcomes.

The aids of automatic machine translation

We could see from our results that although the quality of MT is not very good (because translation error often occurs), the translated contents are still the first language of our participants. As a result, our participants could understand most of the translated contents.

"I think the translation is awful, but due to the fact that I speak Chinese, I could still understand [the translated contents] no matter how the word order was messed up." [P13]

In addition, they could roughly figure out what each paragraph or the entire article was talking about by reading or glancing over the translated contents.

"[I was] able to approximately understand the meaning of the entire article with Chinese." [P11]

Sentence-level separation and mapping using colors

According to our results, sentence-level separation using colors could help remedy the line skipping problem. Although we did not find any statistically significant difference on comprehension level between reading with and without sentence-level mapping, we found from interview that sentence-level separation and mapping could help second-language readers to better and more conveniently analyze the sentence structure. It also benefits the reading resumption problem and the mapping between original sentences and their corresponding translated contents.

"Separated sentences helped me find the beginning and the ending of the sentence more easily. They also helped me analyze the sentence structure with ease. I don't need to select the entire sentence cause it's already done for me." [P7]

"I was able to look one English sentence and one Chinese sentence iteratively since background color made me find corresponding sentences faster." [P4]

"It's easier to find corresponding sentence distinguished by colors. You won't forget where you've just read when switching between English and Chinese." [P1]

Information overload

Our results also suggest that displaying the translation of the reading paragraph is enough for second-language readers, or it may dazzle the readers.

"Unrelated information would not appear [with active translation], so that I could concentrate on what I was looking at." [P6]

DESIGN IMPLICATIONS

Using Multiple Input Sources

Apart from the translation of a paragraph, participants also wanted to know the meaning of particular words sometimes. Due to the limitation of the low-cost eye tracker and the fact that it is exhausting to control all things by eyes, increasing the input sources may be a good choice for detecting the demands of users. For instance, users could also be allowed to highlight the unfamiliar words by using mouse while using eye tracker to determine which paragraph is being read. System can then highlight the corresponding translation of the words in the translated contents, or show the word translation in an independent space in the same page.

Augmenting the Usage of Highlighting, Mapping and Separation by Colors

Inferring from our results, highlighting by colors could help readers focus on what they were looking at, without being distracted by the irrelevant contents. Mapping by colors could help readers find the corresponding item without using indicators such as lines or arrows, which may cover the reading contents. Separation by colors could help identify different parts or categories from crowded elements. As a result, it is reasonable to extend this concept to other parts on webpages or other interfaces. This may benefit the online reading experience.

Additionally, here is an example of footnoted text.[1] (The footnote is created with the "footnote…" command under the "Insert" menu in MS Word). As stated in the footnote, footnotes should rarely be used.

CONCLUSION

In this work, we first conducted a preliminary study in exploration of proper form of MT for second-language readers, and then a experiment was conducted to examine the concept of our proposed interface for supporting second-language readers. From the results, we found that combining active display of MT using eye tracking technique and sentence-level mapping using background colors is the best design to help second-language readers among other combination in our experiment. Our experiments justified that the proposed combination reduced the workload of second-language readers and enhanced their comprehension level and readability without dazzling them.

In future work, we'll enrich the discussion to indicate possible future directions, including logging eye-tracking data for triangulating the results of the current study, running studies on supporting translations of language pairs other than English-Chinese, and multimodal input as a way for triggering active translation support.

REFERENCES

1. Copeland, L., & Gedeon, T. (2013, November). The effect of subject familiarity on comprehension and eye movements during reading. In Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration (pp. 285-288). ACM.
2. Feng, W., Chen, M., & Betke, M. (2014, May). Target reverse crossing: a selection method for camera-based mouse-replacement systems. In Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments (p. 39). ACM.
3. Han, C. H., Yang, C. L., & Wang, H. C. (2014, April). Supporting second language reading with picture note-taking. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems* (pp. 2245-2250). ACM.
4. Hohlfeld, O., Pomp, A., Link, J. Á. B., & Guse, D. (2015, August). On the applicability of computer vision based gaze tracking in mobile scenarios. In Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (pp. 427-434). ACM.
5. Hyrskykari, A., Majaranta, P., Aaltonen, A., & Räihä, K. J. (2000, November). Design issues of iDICT: a gaze-assisted translation aid. In Proceedings of the 2000 symposium on Eye tracking research & applications (pp. 9-14). ACM.
6. Jo, Jaemin, Bohyoung Kim, and Jinwook Seo. 2015. EyeBookmark: Assisting recovery from interruption during reading. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* 2963-2966.
7. Kaomanfen. Retrieved July 17, 2017 from http://www.kmf.com/
8. Kern, Dagmar, Paul Marshall, and Albrecht Schmidt. 2010. Gazemarks: gaze-based visual placeholders to ease attention switching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2093-2102.

9. Kurauchi, A., Feng, W., Joshi, A., Morimoto, C., & Betke, M. (2016, May). EyeSwipe: Dwell-free text entry using gaze paths. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 1952-1956). ACM.

10. Lutteroth, C., Penkar, M., & Weber, G. (2015, November). Gaze vs. mouse: A fast and accurate gaze-only click alternative. In Proceedings of the 28th annual ACM symposium on user interface software & technology (pp. 385-394). ACM.

11. Okoso, Ayano, Kai Kunze, and Koichi Kise. 2014. Implicit gaze based annotations to support second language learning. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication.* 143-146.

12. Penkar, A., Lutteroth, C., & Weber, G. (2013, September). Eyes only: Navigating hypertext with gaze. In 14th International Conference on Human-Computer Interaction (INTERACT) (No. Part II, pp. 153-169). Springer.

13. Project IRIS. Retrieved on Apirl 20, 2017 from http://iris.xcessity.at/

14. Test your English. Cambridge English Language Assessment. Retrieved July 17, 2017 from http://www.cambridgeenglish.org/test-your-english/adult-learners/

15. Yu, Chen-Hsiang, and Robert C. Miller. 2010. Enhancing web page readability for non-native readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2523-2532.