The Effect of Extensive Reading on Students' Ability to Predict Following Text and Increase Reading Speed

Thomas N. Robb
Kyoto Sangyo University
trobb@cc.kyoto-su.ac.jp

Kenichi Kamiya
Osaka Institute of Technology
kamiya@cc.kyoto-su.ac.jp

Abstract

Fluent readers in any language do not focus on each word while reading because much of the text is predictable from the context and syntactic clues. Yet one reason that those who are learning to read a new language must read slowly is that they lack this ability to predict what is coming next. This project aims to demonstrate, through the "Scrolling Cloze" (SC) software developed by the author, that students, through their extensive reading, improve their ability to predict what will come next. The SC presented the students with text containing blanks that scrolled by quickly and then disappeared from view. The students' task was to select the appropriate word for each blank from a panel of word buttons on their screen. Through a pre-test, post-test format, a cohort of English majors at a Japanese university attempted the Scrolling Cloze activity in April, at the beginning of their first year, and then a second time in December. The results indicate that the students' improvement in the number of correct answers correlated highly with the number of words read throughout the year, although not with any other measure such as improvement in their TOEFL scores, or other online practice activities. Further, there does not appear to be any direct relationship between the number of words read in extensive reading and an increase in reading speed over the course of the academic year.

Introduction

There are now numerous studies that demonstrate that Extensive Reading is an effective approach, at least for some populations of learners. See Nation & Waring (2019, 101) for a listing of some of these studies.

The Extensive Reading Foundation's website lists these outcomes of ER, citing one or more studies for each claim:

1) ER increases reading skills, speed, and understanding.
2) ER builds vocabulary.
3) ER expands grammar knowledge.
4) ER develops listening and speaking ability.
5) ER correlates with improvement on tests like the TOEIC.
6) ER increases student enjoyment of reading.

One skill, however, is not mentioned, although it is related to those listed above: the ability to predict what is coming next in the text. Naturally, the more the reader can predict what is coming next, the less attention is required to those words. In fact, readers simply skip over the words that they have guessed from context, and this is one reason that good readers can read more quickly.

Let us look at a practical demonstration in the form of a common puzzle:[i]

> Read the following passage ONCE quickly from start to finish and count the number of f's that you find.

Finished files are the result of years of scientific study combined with the experience of years...

The result is often that native speakers and those very proficient at English will see only three Fs, while slow readers will catch all of them, including those in the three occurrences of the word "of." This happens because the preposition is predicable from the preceding syntax.

Another argument is that native speakers, unless trained to proofread, very often skip over errors simply because they have already assumed what the words were. Thus, some recommended strategies for effective proofing include reading the text backwards, holding the paper upside down, or using a ruler–all of which force the proofreader to slow down and pay attention to words that otherwise might get skipped.

Research using eye-tracking

In eye-tracking research with native speakers, "It is established and unchallenged that function words are fixated on about 35% of the time and content words are fixated on about 85% of the time (Rayner and Duffy, 1988)," as cited in Barrett & Søgaard (2015).

Non-native speakers have difficulty recognizing idioms and therefore read them at a normal reading rate, whereas native speakers comprehend the phrase more quickly and move on with their reading. According to Siyanova-Chanturia, et al. (2011) "Contrary to native speaker results, non-native findings suggest that L2 speakers process idioms at a similar speed to novel phrases. Further, figurative uses are processed more slowly than literal ones. Importantly, the recognition point analysis allows us to establish where non-natives slow down when processing the figurative meaning" (p. 251).

In research using brain wave measurement on various types of cloze items, Ito (2016) found that "L1 speakers predict word form information, but L2 speakers do not" (p.ii). While she says categorically that they do not predict word form information, it is probably closer to the truth that the more proficient the L2 speakers are, the closer they approach the native-speaker norm.

The Research Plan

The overall premise of this research was that students' ability to guess what is coming next in a sentence would improve along with their reading ability reflected by the number of words that they had read in the extensive reading assignments over two 15-week terms.

Thus our main hypotheses were:

1. Can the SC activity measure students' ability to guess "what is coming next" in a text?

2. Is there a relationship between the amount of extensive reading that students do over the course of an academic year and improvement in their ability to guess "what is coming next"?

While a standard cloze text can measure students' ability to guess, we were more interested in their ability to do this quickly and accurately as they are reading. A paper-based cloze would permit them to mull over each item and attempt to guess the word from the larger context, both before and after the blank line. Thus we opted for a "scrolling cloze" (SC) test that would provide only a limited amount of time for them to respond. The number of correct items would then be correlated with the number of words read during the year. Other factors related to language ability and practice would also be examined to the extent possible. (We could not include class grades or in-class performance, amount of self-initiated use of English outside of class, and many other factors that would also surely contribute to improved overall ability in English.)

The number of words read would be taken from their word count on MReader, a program previously developed by the researcher. MReader (http://mreader.org) provides a randomized quiz of 10 questions from a larger question bank for each book that the students have read. The student, as well as the instructor, can then track reading progress. Naturally, the word count assumes that the students have actually read each of the books that they have claimed that they have read, but there is always a modicum of cheating that renders these word counts less than fully accurate.

Design and Development of the Testing Instrument

The basic idea was to have the students fill in blanks in passages that were available for only a limited amount of time. There were two basic options for achieving this: 1) to show a static passage on the screen with blanks for the students to fill in, but with a timer to place pressure on them to respond quickly; or 2) to have the text scroll past them on the screen with the blank appearing on the right side and slowing moving to the left and off-screen, somewhat like digital news scrolling at the bottom of a TV screen ("marquee function").

We opted for the marquee format since it allowed each word to be displayed for exactly the same length of time regardless of the length of the passage and the number of clozed items included.

Having the students type their responses was not an option since this would introduce extraneous variables such as typing speed, the need to accept or reject misspellings, etc. Therefore, we decided to display a set of choices on the screen so that the subjects merely had to click on their answer choice.

Next, we had the choice of supplying a new set of distractors for each cloze item but this would have taken considerable time to develop and refine. The other option was to simply display the correct choices for each clozed word in the text currently being displayed even though some of the items would clearly (to native speakers) not fit. The screen display thus looked like those in Figure 1. Five different colors were used for the word buttons so that the user could match each button set to a specific blank in the scrolling text.

Figure 1 – The items choices for a sample passage.

Selecting the passages to use and the items to cloze within each passage.

To determine the clozed words, we did not employ a strict policy on the word type or number of intermediate words. Instead, we hoped that through piloting of our initial choices we would be able to determine a set of words that correlated well with our initial information about the subjects' ability in English. We attempted to select words that we, as instructors, could assume that students would be able to guess, either from the context or due to a grammatical requirement, such as a specific preposition required by the preceding verb. We refined the items selection, removing words that either were too easy or too difficult to guess, or which had a low correlation with student ability, by selecting another nearby word as a replacement. More pilot trials might have resulted in a better selection of texts and item choices; however, due to time constraints, there was a limit to how many times we could ask students to pilot the passages.

With permission from the publisher, Cengage, we selected passages from a number of their graded readers in the Page Turners series. We piloted seven texts and reduced the number to four that appeared to yield the best results. The piloting was conducted over the span of a year since there were not so many opportunities to have students trial the student versions, and a large number of responses was required to provide useful data for determining the effectiveness of each clozed word. In fact, in the first two years that the experiment was attempted, defects in the selection of the items as well as faults with the administration procedure afforded more opportunities for improving the texts.

In addition to the graded reader passages, the last two texts in the series were created from 4-word NGRAM collocations culled from the Corpus of Contemporary American English, Brigham Young University. Two sets of 10 items were embedded in a simple sentence from which one word was selected as the target. These were added under the assumption that knowledge of collocations in English would be highly correlated with general language proficiency (Namvar, 2012, p.41).

The actual texts used are reproduced in the Appendix.

Visuals of the software

Figures 2 through 5 are screenshots of the SC application at various steps in actual usage.

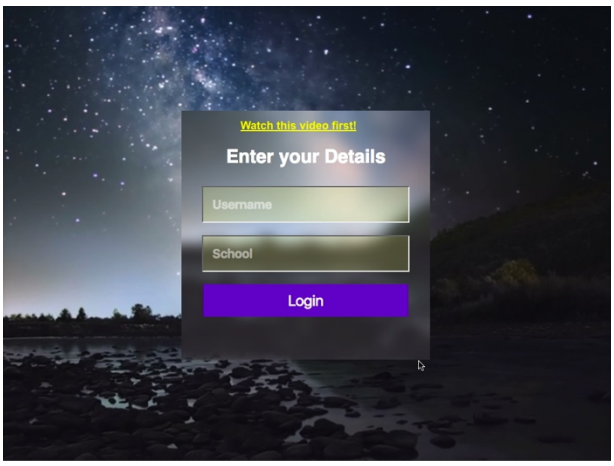

Figure 2 – Login Screen

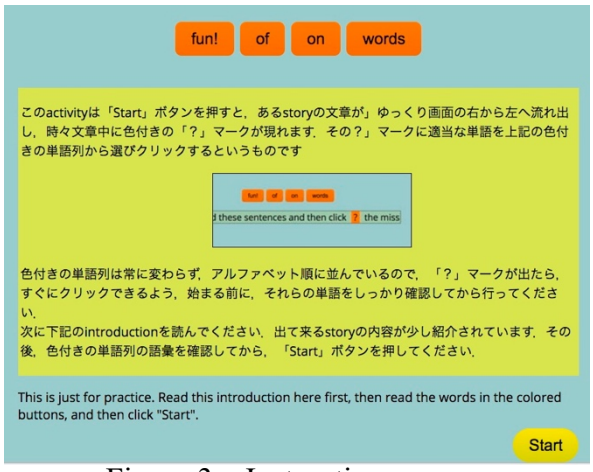

Figure 3 – Instructions

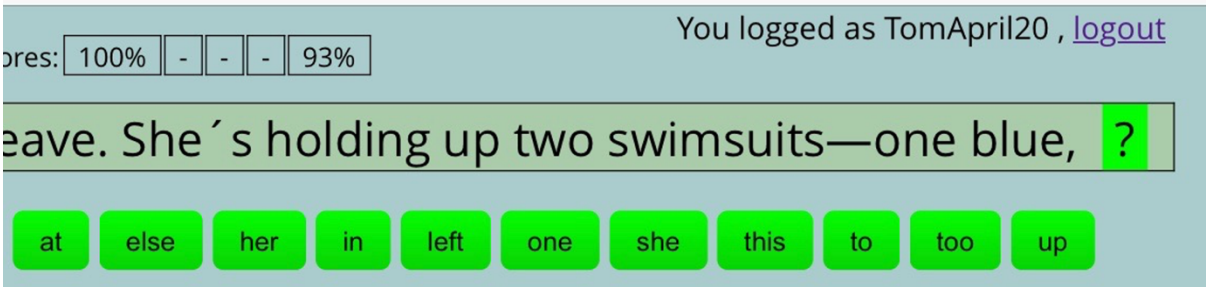

Figure 4 – A text showing a missing word, the "?" with a green background

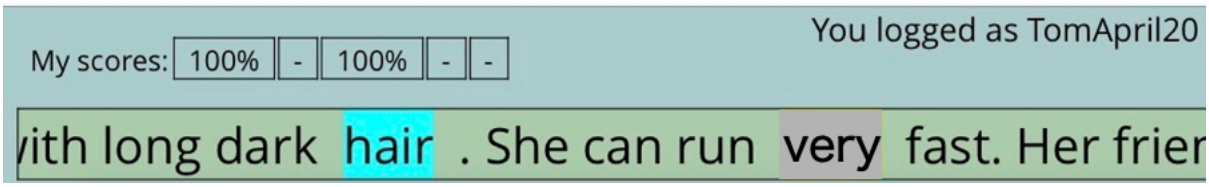

Figure 5 – A text showing a correctly selected word ('hair'), and a corrected error ('very')

Procedure

Once the testing instrument, the "Scrolling Cloze Test" (SCT), was developed, four classes of students of the English Department of Kyoto Sangyo University took the test twice, once in April and again in December, using exactly the same contents. We assumed that a span of six months was sufficient for the students to have forgotten the contents of the first administration. The alternative, to build two equivalent tests, was not an option.

The participants took the test in their "Digital Literacy" classes of which there were four sections with a different instructor for each, all of whom were using the same syllabus and material. The differences in the number of words correct was hypothesized to correlate well with the amount of extensive reading the students did during the academic year from April through December. In actuality, the only results reported here are from the 2018 cohort. Attempts were made in 2016 and 2017, but there were problems with the reliability of the selected cloze words and accidental additional attempts during the term with the SCT contaminated the results. Even in 2018, two of the four cohorts were accidentally allowed access to the SCT during the term, so their data was removed, leaving only 48 valid subjects.

Using the SCT with the students

For each of the four "Joho" (Digitial literacy) classes, the researcher visited the class in the computer room just prior to the start of class. Their instructor introduced me, after which we passed out an explanation on paper with an agreement for them to sign (April administration only). After this, the SCT was demonstrated using the same "Have fun!" passage that the students would use for practice on their own computers. After the introduction to the full class, the students were requested to input their student number and the access code and then start with the "Have fun!" passage (https://www.langconcepts.com/sc5/mp4/HaveFun2.mp4) for practice. There was a link at the end to allow them to repeat the demo, if they felt that they needed more practice. Finally, they were told to start the actual SC passages. The upper right of their screens kept a tally of the percent that they got correct on each passage, although for data crunching purposes, we simply used the number of correct answers over the entire test (excluding the practice passage). While the students were working, the researcher wandered around to help students who were having problems accessing the site and logging in.

Results

Primary Data

1. The number of correct responses to the April and December SC tests. There are "SC_Corr_Apr", "SC_Corr_Dec" as well "SC_Corr-Improv" which is the mathematical difference of the scores.
2. The results of the three sections of the TOEFL – Listening, Structure and Reading. --According to department policy, the students took the shorter "Pre-TOEFL" in April which is set on the same scale as the standard ITP (Institutional Testing Program) version, but with an upper limit of 50 for each subtest instead of 67, with the total score maximum being 500 instead of 667. This shorter version has been deemed suitable since usually no students "top out" in April with a full score of 500, and the test takes less time as well as being less expensive. The full TOEFL is administered in December, towards the end of the academic year.
3. The total number of words read, as recorded by the MReader.org quiz program, which is used to track student reading throughout the year.

Secondary Data

4. Reading speed tests: The entire student body took a reading speed test, internally developed and used continuously for many years. All students took the test in a large room, noting their finishing time from a time display on the front screen, and then proceeded to answer some comprehension questions concerning the text.
5. Word Engine: WE-Correct. An online vocabulary study program. Students are required to obtain 150 correct responses each week.
6. English Central: Videos that the students watched, repeating selected phrases and studying vocabulary from. These activities could be followed by a 20-minute "GoLive" online discussion. Three scores were employed:
    a. VW: Videos Watched. The total number of videos that the student had watched
    b. TR: Total Recordings. How many times the student had recorded the lines when prompted to on the video. This is a measure of active involvement in speaking.
    c. GL: GoLive. The number of times students connected online with a teacher to discuss the videos that they had watched.

Analysis

Research Question 1. Can the SC activity measure students' ability to guess "what is coming next" in a text?

The assumption is that the students' ability to guess the missing word on the SC activity is truly measuring their ability to guess what is coming next from the limited preceding context. This being the case, from Table 1, we can see that there is a highly significant initial correlation between each component of the TOEFL and their scores on the initial SC. From these data we can assume that the SC is indeed measuring some aspect of language ability. Note that there is an insignificant, negative correlation with the total number of words read at the end of the year. If there were a direct correlation between language ability and the amount of reading permitted by their own ability, we would expect some degree of correlation with their SC performance, but there was none. (The full correlation matrix is available in Appendix B.

Table 1. Correlations between the April administration of the SC with initial TOEFL scores.

Pearson's Correlations ▼

Variable		SC_Corr-Apr
1. SC_Corr-Apr	Pearson's r	—
	p-value	—
4. Words	Pearson's r	−0.20
	p-value	0.92
5. TOEFL_L_Apr	Pearson's r	0.55***
	p-value	< .001
6. TOEFL_S_Apr	Pearson's r	0.37**
	p-value	0.00
7. TOEFL_R_Apr	Pearson's r	0.40**
	p-value	0.00

Research Question 2. Is there a relationship between the amount of extensive reading that students do over the course of an academic year and improvement in their ability to guess "what is coming next"?

From Table 2 we can see that there is a highly significant correlation between the words read over the entire academic year and the students' improvement on the SC test. We cannot expect a "high correlation" since there are surely other factors or abilities attained during the school year that moderate the size of the correlation. Note, however, that there is no correlation with any other variable, save for "Listening Improvement" on the TOEFL—another language skill that is also more effective if the listener can guess "what is coming next."

Table 2. Correlations between the December administration of the SC with initial TOEFL scores and total words of Extensive Reading.

Pearson's Correlations

Variable		SC_Corr-Dec	SC_Corr-Improv	Words
2. SC_Corr-Dec	Pearson's r	—		
	p-value	—		
3. SC_Corr-Improv	Pearson's r	0.33*	—	
	p-value	0.01	—	
4. Words	Pearson's r	0.25*	0.44***	—
	p-value	0.04	< .001	—
5. TOEFL_L_Apr	Pearson's r	0.49***	−0.16	0.22
	p-value	< .001	0.87	0.07
6. TOEFL_S_Apr	Pearson's r	0.41**	−0.05	0.17
	p-value	0.00	0.64	0.12
7. TOEFL_R_Apr	Pearson's r	0.52***	0.01	0.30*
	p-value	< .001	0.47	0.02
8. TOEFL_L_Dec	Pearson's r	0.50***	0.03	0.20
	p-value	< .001	0.42	0.09
9. TOEFL_S_Dec	Pearson's r	0.37**	−0.18	0.08
	p-value	0.01	0.89	0.30
10. TOEFL_R_Dec	Pearson's r	0.19	0.19	0.10
	p-value	0.09	0.09	0.25
11. TOEFL-L-Improv	Pearson's r	−0.09	0.26*	−0.07
	p-value	0.74	0.04	0.67
12. TOEFL-S-Improv	Pearson's r	−0.02	−0.12	−0.08
	p-value	0.56	0.79	0.71
13. TOEFL-R-Improv	Pearson's r	−0.33	0.16	−0.20
	p-value	0.99	0.13	0.92

Relationships with other variables

From the data in Table 3, we can observe only one correlation with SC_Corr-Improve, which is a mildly significant correlation with variable "GL"–how many 20-minute sessions the student held with a teacher.

Table 3. Correlations between other variables

Pearson's Correlations

Variable		SC_Corr-Apr	SC_Corr-Dec	SC_Corr-Improv	Words
14. WE_Correct	Pearson's r	3.50e −4	0.16	0.14	0.56***
	p-value	0.50	0.14	0.18	< .001
15. VW	Pearson's r	0.12	0.05	−0.08	0.30*
	p-value	0.21	0.37	0.71	0.02
16. TR	Pearson's r	0.14	0.17	−7.64e −3	0.39**
	p-value	0.18	0.13	0.52	0.00
17. GL	Pearson's r	−0.24	−1.05e −3	0.26*	0.47***
	p-value	0.95	0.50	0.04	< .001

* p < .05, ** p < .01, *** p < .001, one-tailed

Note. All tests one-tailed, for positive correlation

Variable		TOEFL_L_Apr	TOEFL_S_Apr	TOEFL_R_Apr	TOEFL_L_Dec	TOEFL_S_Dec	TOEFL_R_Dec
14. WE_Correct	Pearson's r	0.08	−5.94e −3	0.22	0.19	0.21	0.08
	p-value	0.29	0.52	0.06	0.10	0.08	0.28
15. VW	Pearson's r	0.07	0.02	0.20	0.05	0.14	−2.17e −3
	p-value	0.33	0.45	0.09	0.37	0.16	0.51
16. TR	Pearson's r	0.03	0.20	0.09	0.13	0.11	0.09
	p-value	0.41	0.09	0.26	0.19	0.22	0.27
17. GL	Pearson's r	−0.03	0.02	0.08	0.12	0.02	0.04
	p-value	0.58	0.44	0.30	0.21	0.44	0.40

Variable		TOEFL-L-Improv	TOEFL-S-Improv	TOEFL-R-Improv	WE_Correct	VW	TR	GL
14. WE_Correct	Pearson's r	0.11	0.20	−0.14	—			
	p-value	0.24	0.09	0.83	—			
15. VW	Pearson's r	−0.03	0.11	−0.19	0.57***	—		
	p-value	0.58	0.22	0.91	< .001	—		
16. TR	Pearson's r	0.10	−0.07	−6.21e −3	0.54***	0.77***	—	
	p-value	0.24	0.67	0.52	< .001	< .001	—	
17. GL	Pearson's r	0.18	−3.32e −5	−0.04	0.45***	0.52***	0.59***	—
	p-value	0.12	0.50	0.61	< .001	< .001	< .001	—

Principle Components Analysis

A principle components analysis uses the correlations among the variables to ascertain the degree of commonality in what they are measuring. Those variables that have a high value are then arranged under the same "component." A variable may load on multiple components, as do two of the variables in our analysis in Table 4. The analyst assigns a name to each component based on the main reason for their commonality.

In our analysis, we can observe the following:

Ability: All components of the TOEFL save for "December-Reading." Both of the SC administrations also loaded on this factor.

Study: All components here are measures of actual online study activity, primarily the amount of activity rather than how well students performed.

ER: Only two components, the number of SC correct answers and the amount of extensive reading, had a positive loading on this factor. SC-Corr in April had a negative loading.

Grammar: All measures of grammar loaded on the same, separate factor, most likely because this component of the TOEFL is measuring discrete knowledge more than ability to perform in the language.

Improvement: Improvement in TOEFL Listening and Reading, along with the December Reading score, for which I have no ready explanation.

Table 4. Principle components analysis

Component Loadings	Ability	Study	ER	Grammar	Improvement	
	RC1	RC2	RC3	RC4	RC5	Uniqueness
SC_Corr-Apr	0.55		−0.68			0.14
SC_Corr-Dec	0.77					0.37
SC_Corr-Improv			0.97			0.10
Words			0.62			0.26
TOEFL_L_Apr	0.76					0.25
TOEFL_S_Apr	0.90			−0.54		0.19
TOEFL_R_Apr	0.67					0.25
TOEFL_L_Dec	0.86					0.33
TOEFL_S_Dec	0.49			0.61		0.18
TOEFL_R_Dec					0.79	0.24
TOEFL-L-Improv					0.68	0.47
TOEFL-S-Improv				1.02		0.06
TOEFL-R-Improv					0.67	0.29
WE_Correct		0.69				0.34
VW		0.93				0.20
TR		0.97				0.15
GL		0.72				0.35

Note. Applied rotation method is promax.

Relationship between Reading Speed and performance on the variables under study.

The Department of English administers a reading speed test under tightly controlled conditions at the beginning of the students' first year, and then twice again–at the end of their first and second years. Data from the 2017 and 2018 entrants (N=163) shows an improvement of 41 words per minute on average, and the data correlates highly ($p < .001$) with the December TOEFL scores. Nevertheless, with the small sample of 44 students of the 2018 cohort who participated in the SC research, there were no significant correlations with any of the factors under study. From this we can conclude that there is probably no relationship between improvement in language ability, as measured by the TOEFL, or the SC activity. We can speculate that while reading speed itself did not improve, there was improvement in the students' ability to comprehend what they had read.

Spinoffs of the Scrolling Cloze activity

Although not directly related to the main focus of this research, two other applications for the Scrolling Cloze activity have been developed and are in active use (Robb, 2018).

As a Reading Practice Activity

An interface has been designed so that teachers may input their own texts and use them with their students.

1) As a Reading Placement Test

Teachers may have students use the SC app as a placement test prior to using the http://xreading.com reading software.

Discussion & Limitations

This study was conducted with English majors at a single university. In Japan, as in many other countries, the average language level of the students reflects to a great extent the level of the university that they manage to enter. Thus, if the study were carried out at a different institution, or with students with greater or lesser motivation to study, the results might be different. Furthermore, clearer results might be obtained when the students have read a greater amount for their extensive reading work.

The Scrolling Cloze activity is also limited by the number of passages that can be employed without unduly tiring the students, after which the accuracy and validity of the measurement would slowly decrease.

Conclusion

Despite the limited data that could be utilized, it does indeed appear that extensive reading promotes prediction of what comes next in the text. The software used in this study is available for further studies in other educational contexts.

Acknowledgement

This research was funded by MEXT/JSPS Grant-in-Aid for Scientific Research (Japan), B1-16K01140.

References

Barrett, M. & Søgaard , A. (2015). Reading behavior predicts syntactic categories. *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*.

https://www.aclweb.org/anthology/K15-1038.pdf.

Extensive Reading Foundation (nd). Evidence that ER Works. https://erfoundation.org/wordpress/evidence/.

Itoh, Aine, (2016). Prediction during native and non-native comprehension: The role of mediating factors. Ph.D. Dissertation, The University of Edinburgh, Department of Psychology. https://www.researchgate.net/publication/310748000.

Namvar, F. (2012). The relationship between language proficiency and use of collocation by Iranian EFL students. 3L: *The Southeast Asian Journal of English Language Studies, 18*(3): 41–52. http://journalarticle.ukm.my/5559/1/18_3_5_Firooz.pdf

Nation, I.S.P. & Waring, R. (2019). Teaching Extensive Reading in Another Language, Routledge.

PageTurners, Rob Waring, Series Editor. National Geographic/Cengage Publishers. http://cengagejapan.com/elt/Readers/page/?no=128644549488pkh.

Rayner, K. & Duffy, S. A. (1988). On-line comprehension processes and eye movements in reading, in MacKinnon et al, editors, *Reading research: Advances in theory and practice*, 13–66. Academic Press.

Robb, T. & Kamiya, K. (2017). The Scrolling Cloze activity for measuring reading fluency. https://www.dropbox.com/s/xamyedtuqp188mk/ScrollingCloze-GLoCALL2017-Sept2-Recorded-Edited.pptx?dl=0.

Robb, T. (2018). The Scrolling Cloze for quick assessment of reading level. http://tomrobb.com/CALICO_Robb_Cloze.pptx.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and nonnative speakers. *Second Language Research, 27*(2), 251-272. 10.1177/0267658310382068.

Software & Websites cited

Extensive Reading Foundation – https://erfoundation.org
JASP – https://jasp-stats.org/
MReader – https://mreader.org

Appendix A

Passages employed for the SC tests. Words preceded by a "/" were the selected items. The software culled these words for display in the word buttons and substituted a "?" for them in the displayed scrolling text.

Practice

Read the introduction at the bottom and then click "Start". Then click /on the missing /words . You can see all /of the /words in the colored buttons. The /words move quickly, /so you have to /work fast. Have /fun !

Soccer Crazy – Level 1

Estela Ramos is running hard down the soccer field. She's tall and very /pretty , with long /dark hair . She can /run very fast. Her friend, Katy, is running /down the right side of /the field. She has the ball /at her feet. Katy's a young woman with blonde hair and /a big smile.

Kitchen Love – Level 3

"More cups, Janine!" shouted Gayle King. Not again, thought Janine Cole. It was near the end /of a busy lunchtime /in the college cafeteria, and Janine felt very tired . Janine's manager, Gayle, made her /tired -- she was always shouting at /her . Janine walked to /the table /near the window. As she walked across the room she looked /at the clock /on the wall. Two-thirty. It was almost /time to pick up Amber, her little three-year-old daughter, from the babysitter's house. It cost money, but it was the only /way that Janine could /work .

The Lift – Level 3

The problem with shopping, thinks Catherine, /while standing in Seattle's most expensive department store, is it costs /too much money. She didn't even /want to leave the house /this weekend. It was the middle of the semester, and she didn't have much money /left . She was just /going to stay home and do some homework. That was before /her best friend Hilary /called her up and said, "Hey, Catherine, there's a sale on /at Linden's". So here she is /at Linden's, on Saturday afternoon, a new top in /her bag, wanting to get home before /she buys anything /else . Her friend Hilary isn't ready /to leave. She's holding up two swimsuits--one blue, /one red — and saying "Catherine, I just can't decide."

NGRAM Set 1
These are not a story! Choose the word that you think of first for each of these 10 phrases. Ready?. . . .
1. At the end of the /day . . . .
2. Thank you very much /for the flowers!
3. It's at the top of /the list.
4. It has nothing to do /with him.
5. You don't /have to go.
6. I don't /think that he likes it.
7. It's on the /other side of the city.
8. And now, for the first time /in history.
9. I don't know if he /wants to go.

10. It happened at the turn of the /century .

NGRAM Set 2

These are not a story, either! Choose the word that you think of first. Ready?. . . .

1. If you don't /want to , that's okay.
2. One of the /things that I remember is...
3. I don't know /about that at all
4. Tomorrow is going to /be a beautiful day.
5. I think it's the right thing to /do .
6. Yes, but at the same /time of the day...
7. Sorry, I don't want to /go shopping now.
8. And it turned out to be a lot /of fun.
9. I know a lot of people /who don't like coffee.
10. In my opinion, I don't /think it's a good idea.

Appendix B

In order to make the data readable, the full correlation matrix has been divided into three sections below.

Pearson's Correlations ▼

Variable		SC_Corr-Apr	SC_Corr-Dec	SC_Corr-Improv	Words
1. SC_Corr-Apr	Pearson's r	—			
	p-value	—			
2. SC_Corr-Dec	Pearson's r	0.49***	—		
	p-value	< .001	—		
3. SC_Corr-Improv	Pearson's r	−0.66	0.33*	—	
	p-value	1.00	0.01	—	
4. Words	Pearson's r	−0.20	0.25*	0.44***	—
	p-value	0.92	0.04	< .001	—
5. TOEFL_L_Apr	Pearson's r	0.55***	0.49***	−0.16	0.22
	p-value	< .001	< .001	0.87	0.07
6. TOEFL_S_Apr	Pearson's r	0.37**	0.41**	−0.05	0.17
	p-value	0.00	0.00	0.64	0.12
7. TOEFL_R_Apr	Pearson's r	0.40**	0.52***	0.01	0.30*
	p-value	0.00	< .001	0.47	0.02
8. TOEFL_L_Dec	Pearson's r	0.37**	0.50***	0.03	0.20
	p-value	0.00	< .001	0.42	0.09
9. TOEFL_S_Dec	Pearson's r	0.46***	0.37**	−0.18	0.08
	p-value	< .001	0.01	0.89	0.30
10. TOEFL_R_Dec	Pearson's r	−0.02	0.19	0.19	0.10
	p-value	0.56	0.09	0.09	0.25
11. TOEFL-L-Improv	Pearson's r	−0.31	−0.09	0.26*	−0.07
	p-value	0.98	0.74	0.04	0.67
12. TOEFL-S-Improv	Pearson's r	0.09	−0.02	−0.12	−0.08
	p-value	0.27	0.56	0.79	0.71
13. TOEFL-R-Improv	Pearson's r	−0.41	−0.33	0.16	−0.20
	p-value	1.00	0.99	0.13	0.92
14. WE_Correct	Pearson's r	3.50e −4	0.16	0.14	0.56***
	p-value	0.50	0.14	0.18	< .001
15. VW	Pearson's r	0.12	0.05	−0.08	0.30*
	p-value	0.21	0.37	0.71	0.02
16. TR	Pearson's r	0.14	0.17	−7.64e −3	0.39**
	p-value	0.18	0.13	0.52	0.00
17. GL	Pearson's r	−0.24	−1.05e −3	0.26*	0.47***
	p-value	0.95	0.50	0.04	< .001

* p < .05, ** p < .01, *** p < .001, one-tailed

Note. All tests one-tailed, for positive correlation

**	—	TOEFL_L_Apr	TOEFL_S_Apr	TOEFL_R_Apr	TOEFL_L_Dec	TOEFL_S_Dec	TOEFL_R_Dec
	—						
5. TOEFL_L_Apr		—					
		—					
6. TOEFL_S_Apr		0.49***	—				
		< .001	—				
7. TOEFL_R_Apr		0.42**	0.41**	—			
		0.00	0.00	—			
8. TOEFL_L_Dec		0.69***	0.56***	0.45***	—		
		< .001	< .001	< .001	—		
9. TOEFL_S_Dec		0.42**	0.37**	0.64***	0.43**	—	
		0.00	0.01	< .001	0.00	—	
10. TOEFL_R_Dec		0.01	0.17	0.44***	0.19	0.31*	—
		0.47	0.12	< .001	0.10	0.02	—
11. TOEFL-L-Improv		−0.55	−0.02	−0.05	0.22	−0.07	0.20
		1.00	0.55	0.63	0.07	0.68	0.08
12. TOEFL-S-Improv		−0.05	−0.54	0.23	−0.10	0.59***	0.13
		0.62	1.00	0.06	0.74	< .001	0.19
13. TOEFL-R-Improv		−0.40	−0.24	−0.58	−0.27	−0.35	0.48***
		1.00	0.95	1.00	0.97	0.99	< .001
14. WE_Correct		0.08	−5.94e −3	0.22	0.19	0.21	0.08
		0.29	0.52	0.06	0.10	0.08	0.28
15. VW		0.07	0.02	0.20	0.05	0.14	−2.17e −3
		0.33	0.45	0.09	0.37	0.16	0.51
16. TR		0.03	0.20	0.09	0.13	0.11	0.09
		0.41	0.09	0.26	0.19	0.22	0.27
17. GL		−0.03	0.02	0.08	0.12	0.02	0.04
		0.58	0.44	0.30	0.21	0.44	0.40

	TOEFL-L-Improv	TOEFL-S-Improv	TOEFL-R-Improv	WE_Correct	VW	TR
11. TOEFL-L-Improv	—					
	—					
12. TOEFL-S-Improv	−0.05	—				
	0.63	—				
13. TOEFL-R-Improv	0.23	−0.11	—			
	0.06	0.76	—			
14. WE_Correct	0.11	0.20	−0.14	—		
	0.24	0.09	0.83	—		
15. VW	−0.03	0.11	−0.19	0.57***	—	
	0.58	0.22	0.91	< .001	—	
16. TR	0.10	−0.07	−6.21e −3	0.54***	0.77***	—
	0.24	0.67	0.52	< .001	< .001	—
17. GL	0.18	−3.32e −5	−0.04	0.45***	0.52***	0.59***
	0.12	0.50	0.61	< .001	< .001	< .001

Appendix C

Principal Component Analysis ▼

Chi-squared Test

	Value	df	p
Model	3180.54	61	< .001

Component Loadings ▼

	RC1	RC2	RC3	RC4	RC5	Uniqueness
SC_Corr-Apr	0.55		−0.68			0.14
SC_Corr-Dec	0.77					0.37
SC_Corr-Improv			0.97			0.10
Words			0.62			0.26
TOEFL_L_Apr	0.76					0.25
TOEFL_S_Apr	0.90			−0.54		0.19
TOEFL_R_Apr	0.67					0.25
TOEFL_L_Dec	0.86					0.33
TOEFL_S_Dec	0.49			0.61		0.18
TOEFL_R_Dec					0.79	0.24
TOEFL-L-Improv					0.68	0.47
TOEFL-S-Improv				1.02		0.06
TOEFL-R-Improv					0.67	0.29
WE_Correct		0.69				0.34
VW		0.93				0.20
TR		0.97				0.15
GL		0.72				0.35

Note. Applied rotation method is promax.

Component Characteristics

	Eigenvalue	Proportion var.	Cumulative
RC1	4.50	0.26	0.26
RC2	3.08	0.18	0.45
RC3	2.04	0.12	0.57
RC4	1.82	0.11	0.67
RC5	1.41	0.08	0.76

[i] Searching for "Count how F's" in an Internet search will reveal many variations of this puzzle.