Instructions for Running E-Z Reader 10.3 Simulations

© Erik D. Reichle (March 27, 2022)

1. Introduction

These are instructions for running simulations using the *E-Z Reader* model of eye-movement control in reading. However, before running simulations, I kindly ask that you first read Reichle (2011) or Reichle, Pollatsek, and Rayner (2012) (see References, below) so that you understand how the model works and what it actually does. I also encourage you to have a basic understanding of the Java programming language because this will be necessary to run simulations.

The .zipped file that you have downloaded, *EZReader 10*, contains these instructions and another folder named *EZReader 10.3*. The latter contains the project and source (.java) code for the E-Z Reader model. The project has been compiled and tested using the Apache Netbeans IDE (ver. 12.6) running Java 17.0.0.2. In my experience, Java is robust and you should therefore have no difficulty opening and running the model using whatever version of Java you have installed on your computer. However, if you don't already have Java and/or an IDE, you can easily download them for free using the following links:

Java: https://www.oracle.com/index.html

Apache Netbeans: https://netbeans.apache.org

2. Preliminary Steps

To run the program, you will need two .txt files: (1) a corpus file containing information about the sentences that will be used in the simulation; and (2) a target file specifying the locations of a specific target word in each of the sentences.

In the *EZReader 10.3* folder, the corpus file is named *SRC98-ELP.txt*. This file contains information about the 48 sentences used by Schilling, Rayner, and Chumbley (1998). If you open the file, the top portion will look like Figure 1. Information about the sentences is arranged into four columns indicating each word's: (1) frequency of occurrence (e.g., as tabulated in the English Lexicon Project; Balota et al., 2007); (2) length (i.e., number of letters); (3) cloze predictability (e.g., see Taylor, 1953); and (4) the actual word. Note that the last word in each sentence is indicated by an ampersand symbol (i.e., @). To run simulations using another sentence corpus, the file must be formatted exactly like the example in Figure 1. (Please be careful that the files doesn't include control characters because this can be problematic for the program.)

SRC98-ELP.txt

```
1         6         0.00      Margie
34884     5         0.00      moved
563310    4         0.20      into
297422    3         0.25      her
709084    3         0.65      new
1368      9         0.75      apartment
1977887   2         0.00      at
23099033  3         0.60      the
227504    3         0.10      end
11003606  2         0.95      of
23099033  3         1.00      the
34481     7         0.10      summer.@
23099033  3         0.00      The
4605      9         0.00      principal
15526     10        0.00      introduced
23099033  3         0.80      the
709084    3         0.35      new
74841     9         0.00      president
11003606  2         0.55      of
23099033  3         1.00      the
10114     6         0.00      junior
94260     6         0.70      class.@
90590     4         0.00      None
```

Figure 1. Example corpus file.

There is also a target file in *EZReader 10.3* folder named *SRC98Targets.txt*. This file contains a single column of numbers indicating the within-sentence ordinal positions of specific target words in the sentences. In the Schilling et al. (1998) sentences, these were the high- and low-frequency target words that were of interest in the experiment. If you open the target file, the top part will look like Figure 2. There, the first number, 5, refers to the word “apartment” in the first sentence of Figure 1. Note that “apartment” is actually the sixth word in the sentence, not the fifth. As per standard Java conventions, an array of *N* items is indexed using the numbers 0 to *N*-1. For example, a target word that is the seventh word of a 12-word sentence would be word 6 with the indices for the first and last words being 0 and 11, respectively. Finally, if using another sentence corpus, it is not necessary to have designated target words, but it is necessary to have a target file. Such a file can be dummy coded using a single column of *N*-1 zeros, with *N* equaling the number of sentences in the corpus.

SRC98Targets.txt

```
5
5
3
9
4
5
6
7
1
```

Figure 2. Example target file.

3. Running Simulations

To run E-Z Reader simulations, first open your IDE and then open the project, *EZReader 10.3*. Then open the class containing the main method, *EZReader10.java*. The top part of this class should look like Figure 3.

```java
// I/O file names:
static String corpusFile = "SRC98-ELP"; // name of sentence corpus file
static String outputFile = "SimulationResults"; // name of output file
static String targetFile = "SRC98Targets"; // name of word target file

// Model free parameters:
static double A = 25.0;
static double Alpha1 = 127.0;
static double Alpha2 = 8.0;
static double Alpha3 = 59.0;
static double Delta = 1.21;
static double Epsilon = 1.05;
static double Eta1 = 0.5;
static double Eta2 = 0.1;
static double I = 50.0;
static double ITarget = 50.0;
static double Lambda = 0.25;
static double M1 = 150.0;
static double M2 = 25.0;
static double Omega1 = 6.0;
static double Omega2 = 3.0;
static double pF = 0.01;
static double pFTarget = 0.01;
static double Psi = 7.0;
static double S = 25.0;
static double SigmaGamma = 20.0;
static double V = 50.0;
static double Xi = 0.5;

// Fixed simulation parameters:
static int maxLength = 16; // = # letters + space left of word
static int maxSentenceLength = 50; // maximum # words / sentence
static int NFreqClass = 8;
static int NSentences = 48;
static int NSubjects = 1000;

// Display toggles:
static boolean displayClassMeans = false; // frequency-class means & RMSD
static boolean displayCorpus = false; // word IVs
static boolean displayDists = false; // IOVP distributions, etc.
static boolean displayMeans = true; // mean word DVs
static boolean displayParameters = false; // parameter values
static boolean displayTrace = false; // trace of individual fixations

// Simulation toggles:
static boolean includeRegressions = false; // include regressions?
```

Figure 3. The EZReader10.java class.

Starting at the top of Figure 3, *EZReader10.java* specifies the names and values of variables corresponding to the input/output (I/O) files that are used in running the simulations, E-Z Reader's free parameters, several (fixed) simulation parameters, and several Boolean toggles that control what simulation results are written to the output file and whether or not trials containing inter-word regressions are included in the simulation results. More specifically:

(i) *corpusFile* – This is the name of the sentence corpus file used in the simulation (e.g., *SRC98-ELP.txt*). Change the file name if you want to use a different sentence corpus.

(ii) *outputFile* – This is the name of the file where the simulation results are written. The file does not have to be provided but will instead be created by the program.

(iii) *targetFile* – This is the name of the file containing the target words (e.g., *SRC98Targets.txt*). Change the file name if you want to use a different sentence corpus.
(iv) *A, Alpha1, Alpha2, … Xi* – These are the default values of E-Z Reader's free parameters. These values allow the model to generate eye movements that closely resemble those reported by Schilling et al. (1998).

(v) *maxLength* – This is the maximum word length (i.e., maximum number of letters) in the simulation; it should be set equal to a value that accommodates the longest word in the corpus file.

(vi) *maxSentenceLength* – This is the maximum number of words in any of the sentences in the corpus file.

(vii) *NFreqClass* – This is the number of frequency classes using the ELP word-frequency norms (Balota et al., 2007) and bins defined by $\log_{10}$(frequency).

(viii) *NSentences* – This is the actual number of sentences in the corpus file.

(ix) *NSubjects* – This is the number of statistical subjects used in the simulation.

(x) *displayClassMeans, displayCorpus, … displayTrace* – Setting these Boolean toggles to "true" will cause the program to store different simulation results to the output file. More information about these options is provided in the next section of this document.

(xi) *includeRegressions* – By setting this Boolean toggle to "true," simulated trials containing inter-word regressions will be included in the final simulation output. Please note that the default value of this toggle is "false," with only those trials comprised of first-pass fixations being used to generate the simulation results.

After you have provided the names of the corpus and target files, set the parameters values, and indicated which types of output to store in the output file (by setting the values of the appropriate toggles to "true"), you can start the simulation by clicking on the "run" button in your IDE. (In Apache Netbeans, clicking on the green arrow on the top menu bar will start the program.) The output file should then appear in the program folder, *EZReader 10.3*. Examples and explanations of simulation output are provided below. On a reasonably fast computer, a simulation using 1,000 statistical subjects and the Schilling et al. (1998) sentences should require only a few seconds to complete.

4. Simulation Output

As mentioned in the previous section of this document, the display toggles determine the type of simulation results that will be written to the output file. The possible types of output are as follows:

(i) *displayClassMeans* – Setting this toggle to "true" will store the mean observed and predicted values (across 8 word-frequency classes, *f*) of 6 eye-movement measures: (1) first-fixation duration (*FFD*); (2) single-fixation durations (*SFD*); (3) gaze duration (*GD*); (4) probably of making a single fixation (*Pr1*); (5) probability of making two or more fixations (*Pr2*); and (6) probably of skipping (*PrS*). An example of such output is shown in Figure 4. Within each pair of values, the observed value is displayed first and the simulated value is displayed second. For example, across all of the words in the first frequency class (i.e., *f* = 1), the mean observed and simulated gaze durations are 336 ms and 316 ms, respectively.

Finally, at the bottom of the file are two other values. The first is the *RMSD* between the observed and simulated means (see Appendix of Reichle, Pollatsek, Fisher, & Rayner, 1998), which provides an measure of the simulation's goodness of fit to the observed data. (Smaller *RMSD* values indicate a better fit.) The second is the mean proportion of trials that resulted in inter-word regressions and that were therefore excluded from the analyses.

```
SimResults.txt
f   FFD      SFD      GD       Pr1        Pr2        PrS
-   -------  -------  -------  ---------  ---------  ---------
1   274 259  281 277  336 316  0.73 0.71  0.20 0.20  0.07 0.09
2   237 247  257 262  284 322  0.66 0.64  0.18 0.29  0.17 0.06
3   243 240  253 252  304 287  0.70 0.74  0.25 0.17  0.05 0.09
4   246 237  260 245  301 268  0.67 0.78  0.22 0.12  0.11 0.11
5   238 234  246 241  266 264  0.71 0.74  0.13 0.11  0.16 0.15
6   242 223  246 225  260 242  0.58 0.66  0.05 0.07  0.37 0.27
7   224 217  228 218  235 227  0.37 0.54  0.03 0.03  0.60 0.43
8   205 205  204 205  209 214  0.31 0.47  0.01 0.02  0.68 0.51

RMSD 0.3224 PrReg 0.2544
```

Figure 4. Example output using *displayClassMeans*.

(ii) *displayCorpus* – Setting this toggle to "true" will store the independent variables that are generated for each of word in the sentence corpus. An example of this output is shown in Figure 5. The output is organized by sentences (from the top of the file, starting with Sentence 0) and by words (each row corresponds to a word). The columns thus code the following information about each word's: (1) frequency; (2) cloze predictability; (3) length; (4) cumulative character position of the left edge of the blank space preceding the word; (5) center of word, or *optimal viewing position* (*OVP*; see Vitu, McConkie, Kerr, & O'Regan, 2001); (6) right edge of the word's final letter; and (7) the actual word. For example, the word "Margie" is low frequency (= 1), completely unpredictable (= 0.00), and 6 letters in length. Because the word is the first in the sentence, the beginning of the blank space to the left of the word is located as position 0 and the rightmost edge of the final letter is located at position 7. The OVP for "Margie" is located between the letters "r" and "g," at position 4. Finally, words that are designated targets are marked by asterisk symbols (i.e., *) to make them easier to identify (e.g., for the purpose of analyses). It is useful to check the corpus information prior to running simulations to check that the program is reading the corpus file correctly. (Note that punctuation is included in the word length; e.g., the length of "summer" is 7 because the period ending the sentence is included with the word.)

```
SimResults.txt
Sentence  0
           1 0.00  6   0  4.0   7 Margie
       34884 0.00  5   7 10.5  13 moved
      563310 0.20  4  13 16.0  18 into
      297422 0.25  3  18 20.5  22 her
      709084 0.65  3  22 24.5  26 new
        1368 0.75  9  26 31.5  36 apartment *
     1977887 0.00  2  36 38.0  39 at
    23099033 0.60  3  39 41.5  43 the
      227504 0.10  3  43 45.5  47 end
    11003606 0.95  2  47 49.0  50 of
    23099033 1.00  3  50 52.5  54 the
       34481 0.10  7  54 58.5  62 summer.
Sentence  1
    23099033 0.00  3   0  2.5   4 The
        4605 0.00  9   4  9.5  14 principal
       15526 0.00 10  14 20.0  25 introduced
    23099033 0.80  3  25 27.5  29 the
      709084 0.35  3  29 31.5  33 new
       74841 0.00  9  33 38.5  43 president *
    11003606 0.55  2  43 45.0  46 of
```

Figure 5. Example output using *displayCorpus*.

(iii) *displayDists* – Setting this toggle to "true" will store three kinds of distributions for words of each length: (1) first-fixation landing-site distributions; (2) refixation-probability distributions; and (3) IOVP distributions. The first distributions correspond to the proportions of initial fixations on each letter within a word. For example, for 1-letter words, 0.43 of the initial fixations are on the blank space preceding the word, while the remaining 0.57 land on the word itself. The second distributions show the proportion of initial fixations that were followed by a refixation as a function of the initial fixation's location. For example, for the 5-letter words, initial fixations at the first 3 locations were followed by refixations on 0.39, 0.20, and 0.06 of the simulated trials, respectively. Finally, the last distributions show the mean single-fixation durations as a function of their locations. For example, for 2-letter words, the mean single-fixation duration was 217 ms when the fixation was on the first letter of the word. (Note that there are missing values because the sentences only contained one word that longer than 12 letters and it was excluded from the analyses because it was a sentence-final word.)

SimResults.txt

```
First-fixation landing-site distributions:
 1-letter:  0.43 0.57
 2-letter:  0.24 0.34 0.42
 3-letter:  0.19 0.21 0.31 0.30
 4-letter:  0.10 0.13 0.25 0.31 0.20
 5-letter:  0.09 0.10 0.18 0.28 0.24 0.12
 6-letter:  0.08 0.08 0.13 0.23 0.26 0.16 0.06
 7-letter:  0.10 0.07 0.10 0.18 0.23 0.18 0.09 0.05
 8-letter:  0.09 0.05 0.08 0.13 0.21 0.20 0.13 0.06 0.04
 9-letter:  0.06 0.04 0.04 0.10 0.20 0.24 0.17 0.08 0.04 0.02
10-letter:  0.09 0.06 0.05 0.08 0.15 0.21 0.18 0.10 0.04 0.02 0.01
11-letter:  0.05 0.04 0.04 0.06 0.14 0.19 0.18 0.12 0.08 0.04 0.03 0.03
12-letter:  0.09 0.02 0.02 0.05 0.08 0.15 0.18 0.15 0.08 0.05 0.04 0.06 0.03
13-letter:  0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
14-letter:  0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
15-letter:  0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00

Refixation probability distributions:
 1-letter:  0.01 0.01
 2-letter:  0.05 0.01 0.03
 3-letter:  0.15 0.04 0.02 0.05
 4-letter:  0.29 0.12 0.03 0.04 0.08
 5-letter:  0.39 0.20 0.06 0.02 0.05 0.10
 6-letter:  0.50 0.29 0.12 0.04 0.04 0.08 0.10
 7-letter:  0.59 0.44 0.23 0.10 0.04 0.07 0.12 0.12
 8-letter:  0.60 0.54 0.27 0.13 0.05 0.04 0.08 0.07 0.11
 9-letter:  0.55 0.51 0.31 0.19 0.09 0.04 0.08 0.12 0.13 0.10
10-letter:  0.66 0.76 0.44 0.26 0.17 0.07 0.04 0.07 0.13 0.19 0.08
11-letter:  0.65 0.54 0.40 0.28 0.21 0.09 0.04 0.07 0.13 0.11 0.05 0.10
12-letter:  0.76 0.66 0.37 0.38 0.30 0.21 0.09 0.12 0.12 0.09 0.08 0.04 0.03
13-letter:  0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
14-letter:  0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
15-letter:  0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00

IOVP distributions:
 1-letter:   219  218
 2-letter:   217  217  216
 3-letter:   213  210  212  209
 4-letter:   220  224  229  232  223
 5-letter:   218  228  237  241  235  225
 6-letter:   219  235  245  254  251  240  227
 7-letter:   217  235  250  258  264  251  236  218
 8-letter:   224  228  234  245  249  247  238  221  212
 9-letter:   217  220  226  244  252  258  247  232  216  204
10-letter:   250  217  231  249  257  271  271  253  236  208  225
11-letter:   225  216  218  246  272  271  287  265  251  215  204  196
12-letter:   228  230  199  205  206  205  218  213  210  207  212  195  211
13-letter:     0    0    0    0    0    0    0    0    0    0    0    0    0    0
14-letter:     0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
15-letter:     0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
```

Figure 6. Example output using *displayDists*.

(iv) *displayMeans* – This is the most useful output. Setting this toggle to "true" will store a number of means for each word in the sentence corpus, as shown in Figures 7-10. The beginning of the file (Figure 7) shows eight word-based means: (1) single-fixation duration

(*SFD*); (2) first-fixation duration (*FFD*); (3) gaze duration (*GD*); (4) probability of fixating (*PrF*); (5) probability of fixating once (*Pr1*); (6) probability of fixating two or more times (*Pr2*); (7) probability of skipping (*PrS*); and (8) the word itself. Target words are marked by asterisks. Note that the first and last words in each sentence are not included (per convention) because processing starts and ends abruptly on these words. For example, in the first sentence, the words "Margie" and "summer." are not included in the output.

```
SimResults.txt

Word-based means:
SFD 265  FFD 263  GD 267  TT 267  PrF 0.98 Pr1 0.96 Pr2 0.02 PrS 0.02 moved
SFD 227  FFD 227  GD 233  TT 233  PrF 0.83 Pr1 0.80 Pr2 0.02 PrS 0.17 into
SFD 231  FFD 230  GD 237  TT 237  PrF 0.61 Pr1 0.59 Pr2 0.02 PrS 0.39 her
SFD 224  FFD 223  GD 236  TT 236  PrF 0.53 Pr1 0.50 Pr2 0.03 PrS 0.47 new
SFD 221  FFD 211  GD 276  TT 276  PrF 0.82 Pr1 0.60 Pr2 0.21 PrS 0.18 apartment *
SFD 271  FFD 271  GD 272  TT 272  PrF 0.53 Pr1 0.52 Pr2 0.00 PrS 0.47 at
SFD 223  FFD 221  GD 245  TT 245  PrF 0.56 Pr1 0.49 Pr2 0.07 PrS 0.44 the
SFD 256  FFD 254  GD 267  TT 267  PrF 0.69 Pr1 0.65 Pr2 0.04 PrS 0.31 end
SFD 210  FFD 210  GD 213  TT 213  PrF 0.45 Pr1 0.44 Pr2 0.01 PrS 0.55 of
SFD 179  FFD 180  GD 184  TT 184  PrF 0.60 Pr1 0.59 Pr2 0.02 PrS 0.40 the
SFD 221  FFD 217  GD 233  TT 233  PrF 0.98 Pr1 0.93 Pr2 0.05 PrS 0.02 principal
SFD 270  FFD 260  GD 285  TT 285  PrF 0.95 Pr1 0.84 Pr2 0.11 PrS 0.05 introduced
SFD 206  FFD 206  GD 212  TT 212  PrF 0.56 Pr1 0.54 Pr2 0.02 PrS 0.44 the
SFD 224  FFD 222  GD 229  TT 229  PrF 0.52 Pr1 0.50 Pr2 0.02 PrS 0.48 new
SFD 248  FFD 233  GD 296  TT 296  PrF 0.96 Pr1 0.70 Pr2 0.26 PrS 0.04 president *
SFD 218  FFD 218  GD 218  TT 218  PrF 0.46 Pr1 0.46 Pr2 0.00 PrS 0.54 of
SFD 193  FFD 193  GD 200  TT 200  PrF 0.46 Pr1 0.44 Pr2 0.02 PrS 0.54 the
SFD 248  FFD 241  GD 258  TT 258  PrF 0.84 Pr1 0.77 Pr2 0.07 PrS 0.16 junior
SFD 195  FFD 195  GD 195  TT 195  PrF 0.31 Pr1 0.31 Pr2 0.00 PrS 0.69 of
SFD 183  FFD 182  GD 191  TT 191  PrF 0.59 Pr1 0.56 Pr2 0.03 PrS 0.41 the
SFD 229  FFD 221  GD 251  TT 251  PrF 0.79 Pr1 0.70 Pr2 0.09 PrS 0.21 students *
SFD 260  FFD 247  GD 289  TT 289  PrF 0.89 Pr1 0.75 Pr2 0.15 PrS 0.11 wanted
```

Figure 7. Example output using *displayMeans*. The first shows the word-based means for each word in the sentence corpus.

The next part of the output file (Figure 8) shows the proportions of first-fixations on each possible letter position within a word, with the first "letter" position corresponding to the blank space preceding the word. For example, for the word "moved," 0.01, 0.07, and 0.24 of the simulated first fixations were located on the first, second, and third letter positions, respectively.

```
SimResults.txt
SFD 253  FFD 245  GD 270  TT 270  PrF 0.88 Pr1 0.80 Pr2 0.09 PrS 0.12 inflamed
SFD 230  FFD 229  GD 236  TT 236  PrF 0.74 Pr1 0.72 Pr2 0.02 PrS 0.26 and
SFD 251  FFD 243  GD 280  TT 280  PrF 0.89 Pr1 0.76 Pr2 0.13 PrS 0.11 sore,
SFD 176  FFD 177  GD 184  TT 184  PrF 0.34 Pr1 0.33 Pr2 0.01 PrS 0.66 he
SFD 268  FFD 249  GD 304  TT 304  PrF 0.91 Pr1 0.71 Pr2 0.20 PrS 0.09 visited
SFD 196  FFD 195  GD 213  TT 213  PrF 0.57 Pr1 0.52 Pr2 0.05 PrS 0.43 the
SFD 226  FFD 225  GD 239  TT 239  PrF 0.61 Pr1 0.57 Pr2 0.04 PrS 0.39 eye

First-fixation landing-site distributions:
0.01 0.07 0.24 0.36 0.23 0.09  moved
0.04 0.11 0.27 0.33 0.25  into
0.15 0.15 0.25 0.45  her
0.27 0.24 0.24 0.25  new
0.14 0.07 0.07 0.12 0.18 0.14 0.11 0.07 0.04 0.06  apartment *
0.28 0.39 0.33  at
0.23 0.19 0.28 0.30  the
0.24 0.18 0.29 0.29  end
0.35 0.30 0.35  of
0.31 0.26 0.23 0.21  the
0.00 0.00 0.01 0.10 0.24 0.30 0.22 0.10 0.02 0.00  principal
0.00 0.02 0.07 0.14 0.19 0.23 0.16 0.10 0.04 0.02 0.02  introduced
0.16 0.25 0.31 0.28  the
0.25 0.22 0.26 0.27  new
0.10 0.04 0.07 0.12 0.17 0.23 0.15 0.08 0.02 0.02  president *
```

Figure 8. Example output using *displayMeans*. The figure shows the first-fixation landing-site distributions for each word in the sentence corpus.

The next part of the output file (Figure 9) shows the proportions of initial fixations at each possible location that were followed by refixations, with the first "letter" position again corresponding to the blank space preceding the word. For example, for the word “new,” initial fixations on the first three positions were followed by refixations on 0.14, 0.04, and 0 of the simulated trials.

```
SimResults.txt
0.10 0.07 0.17 0.28 0.25 0.12  sore,
0.18 0.35 0.48  he
0.18 0.08 0.07 0.16 0.18 0.18 0.08 0.06  visited
0.16 0.27 0.33 0.24  the
0.30 0.20 0.25 0.25  eye

Refixation probability distributions:
0.00 0.06 0.01 0.00 0.03 0.05  moved
0.28 0.00 0.00 0.01 0.05  into
0.17 0.02 0.00 0.01  her
0.14 0.04 0.00 0.07  new
0.46 0.55 0.46 0.45 0.14 0.07 0.19 0.10 0.17 0.09  apartment *
0.02 0.00 0.01  at
0.13 0.10 0.14 0.12  the
0.12 0.02 0.04 0.05  end
0.03 0.01 0.01  of
0.04 0.02 0.03 0.03  the
0.00 0.00 0.09 0.07 0.06 0.00 0.05 0.14 0.00 0.00  principal
0.33 0.53 0.35 0.14 0.13 0.07 0.02 0.09 0.06 0.14 0.05  introduced
0.08 0.03 0.01 0.02  the
0.06 0.01 0.01 0.07  new
0.76 0.71 0.42 0.30 0.19 0.07 0.18 0.10 0.25 0.24  president *
0.00 0.00 0.00  of
0.04 0.01 0.00 0.13  the
0.43 0.17 0.10 0.01 0.01 0.03 0.11  junior
0.00 0.00 0.00  of
```

Figure 9. Example output using *displayMeans.* The figure shows the refixation-probability distribution for each word in the sentence corpus.

The final part of the output file (Figure 10) shows the mean single-fixation durations for single fixations at each possible within-word position, with the first “letter” position again being the blank space preceding the word. For example, for the word “moved,” the single-fixation durations on positions 1-3 were 253 ms, 261 ms, and 271 ms, respectively. (The name “IOVP distributions” reflects the fact that—perhaps counter to intuition—single fixations near the OVP tend to be longer in duration than those near either end of the word; see Vitu et al., 2001.)

```
SimResults.txt
0.94 0.00 0.05 0.02 0.01 0.09 0.11  retina *
0.66 0.50 0.11 0.05 0.03 0.04 0.26  became
0.69 0.25 0.25 0.11 0.06 0.03 0.10 0.12 0.26  inflamed
0.07 0.02 0.03 0.04  and
0.62 0.36 0.10 0.04 0.04 0.06  sore,
0.11 0.00 0.07  he
0.63 0.46 0.23 0.09 0.04 0.05 0.16 0.14  visited
0.21 0.10 0.01 0.07  the
0.18 0.05 0.01 0.02  eye

IOVP distributions:
 253  261  271  275  266  241  moved
 221  224  236  234  227  into
 224  217  233  234  her
 217  211  223  210  new
 244  220  208  219  218  221  232  220  224  225  apartment *
 259  269  257  at
 234  208  226  225  the
 224  249  254  249  end
 213  211  209  of
 194  183  173  170  the
   0  159  160  212  221  227  214  203  185    0  principal
 250  216  225  238  265  292  282  261  234  233  192  introduced
 190  212  205  210  the
```

Figure 10. Example output using *displayMeans*. The figure shows the IOVP distributions for each word in the sentence corpus.

(v) *displayParameters* – Setting this toggle to "true" stores the parameter values at the top of the output file. This information is useful so I recommend that the toggle be set equal to "true" by default. (This output can be stored in conjunction with other output, as Figure 11 shows.)

SimulationResults.txt

```
Parameter values:
A = 25.00
Alpha1 = 127.00
Alpha2 = 8.00
Alpha3 = 59.00
Delta = 1.21
Epsilon = 1.05
Eta1 = 0.50
Eta2 = 0.10
I = 50.00
ITarget = 50.00
Lambda = 0.25
M1 = 150.00
M2 = 25.00
Omega1 = 6.00
Omega2 = 3.00
pF = 0.01
pFTarget = 0.01
Psi = 7.00
S = 25.00
SigmaGamma = 20.00
V = 50.00
Xi = 0.50

NSubjects = 1000


Word-based means:
SFD 264  FFD 262  GD 268  TT 268  PrF 0.98 Pr1 0.96 Pr2 0.02 PrS 0.02 moved
SFD 234  FFD 231  GD 243  TT 243  PrF 0.85 Pr1 0.80 Pr2 0.05 PrS 0.15 into
SFD 229  FFD 228  GD 240  TT 240  PrF 0.57 Pr1 0.54 Pr2 0.03 PrS 0.43 her
SFD 222  FFD 222  GD 246  TT 246  PrF 0.56 Pr1 0.50 Pr2 0.07 PrS 0.44 new
SFD 225  FFD 214  GD 276  TT 276  PrF 0.79 Pr1 0.59 Pr2 0.20 PrS 0.21 apartment *
SFD 264  FFD 263  GD 267  TT 267  PrF 0.51 Pr1 0.50 Pr2 0.01 PrS 0.49 at
```

Figure 11. Example output using *displayParameters*. (In this example, the *displayMeans* toggle was also set equal to "true.")

(vi) *displayTrace* – Setting this toggle to "true" will store a simple record of each fixation duration (*dur*) position (*pos*), along with its corresponding word. Figure 12 shows an example. *Warning*: Because simulations generate large number of fixations, only generate this output using a small number of simulated subjects (i.e., *NSubjects* < 10).

SimRes

```
dur: 337 pos:  4.0 word:  0  Margie
dur: 193 pos: 10.9 word:  1  moved
dur: 241 pos: 16.3 word:  2  into
dur: 277 pos: 22.3 word:  4  new
dur: 157 pos: 23.6 word:  4  new
dur: 292 pos: 32.6 word:  5  apartment
dur: 138 pos: 37.1 word:  6  at
dur: 139 pos: 44.5 word:  8  end
dur: 287 pos: 47.2 word:  9  of
dur: 204 pos: 53.9 word: 10  the
dur: 244 pos:  2.5 word:  0  The
dur: 240 pos:  9.0 word:  1  principal
dur: 178 pos: 18.7 word:  2  introduced
dur: 238 pos: 26.9 word:  3  the
```

Figure 13. Example output using *displayTrace*.

5. Final Notes

I've done my best to make the code as easy to understand and use as possible. That being said, I also recognize that it's still a fairly complicated program, so if you run into any problems, please let me know and I'll do my best to help you. Good luck!

6. References

Balota, D., Yap, M. J., Hutchinson, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treisman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445-459.

Reichle, E. D. (2011). Serial attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford handbook on eye movements* (pp. 767-786). Oxford, UK: Oxford University Press.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye-movement control in reading. *Psychological Review*, *105*, 125-157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in non-reading tasks: A unified framework for understanding the eye-mind link. *Psychological Review*, *119*, 155-185.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*, 1270-1281.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, *30*, 415-433.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research*, *41*, 3513-3533.

Other useful references for learning about eye movements in reading include:

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*, 1457-1506.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *The psychology of reading, 2nd Ed*. New York, NY, USA: Psychology Press.

Reichle, E. D. (2021). *Computational models of reading: A handbook*. Oxford, UK: Oxford University Press.