Counting Letters

2022-11-09

Get ready

We'll do an experiment on ourselves. The point of the experiment needs to remain obscure until after the data is collected, so as to not bias the data collection process.

You will receive a printed page with the last paragraph of Darwin's Origin of Species.

- Please read through this paragraph, and circle every letter "t".
- *Please proceed at a normal reading speed. If you ever realize that you missed a "t" in a previous word, do not retrace your steps to encircle the "t".*
- You are not expected to get every "t", so don't slow down your reading to get the "t"s.

Activity

The point of this exercise is to collect data on whether our brains perceive words merely as a collection of letters or if sometimes our brains process words as entities. The logic of this test is that, if words are perceived as units, rather than as collections of letters, then this should be especially true for common words. Hence we will look at the errors made in scanning this paragraph, and ask whether we are more (or less) likely to miss finding a "t" when it is part of a common word.

Compare your results to the answer key that marks all the instances of the letter "t". Note that the answer key marks all "t"s in red, but it also puts boxes around some words. The boxes are drawn around all instances of the letter "t" occurring in common words. "Common" is defined here as among the top-twenty words in terms of frequency of use in English; of these six contain one or more "t"s: the, to, it, that, with, and at. In this passage there are 94 "t"s, and 29 are in common words.

Count how many mistakes you made finding "t"s in common words and in less common words.

Use the appropriate test to ask whether the commonness of a word affects your probability of noticing the "t"s in it. You can create the table needed for the test by a command like:

```
tCountTable <- data.frame(Common = c (25, 4), Uncommon = c(55, 10), row.names = c("Found", "Not found"))
```

Replace 25 and 4 with the number of t's in common words that you found or didn't find, respectively. Replace 55 and 10 with the numbers of t's you found or didn't find in uncommon words.