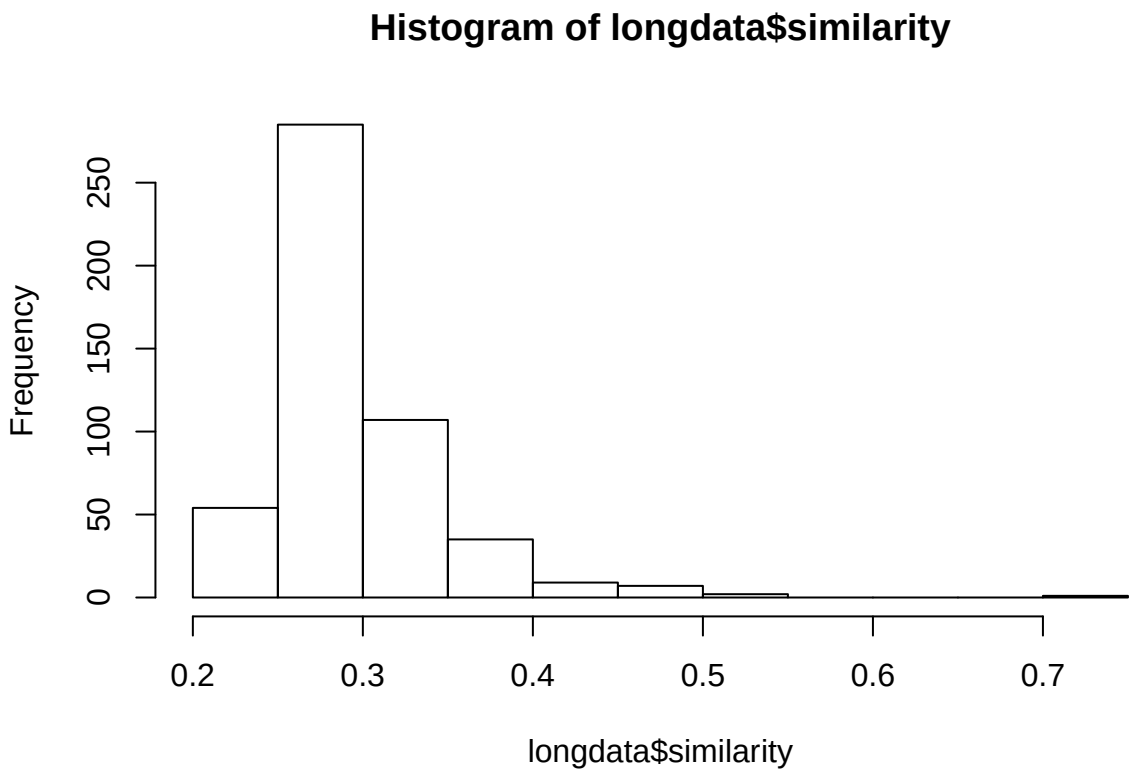
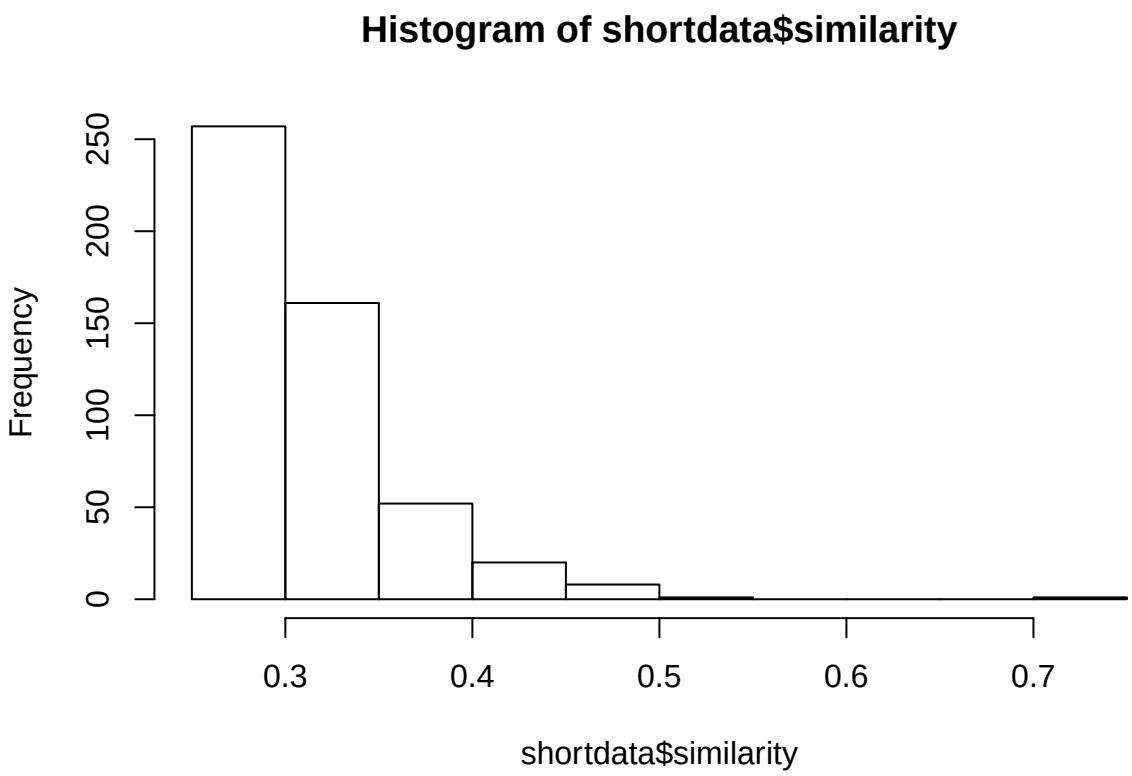
spacetimeadultanalysis

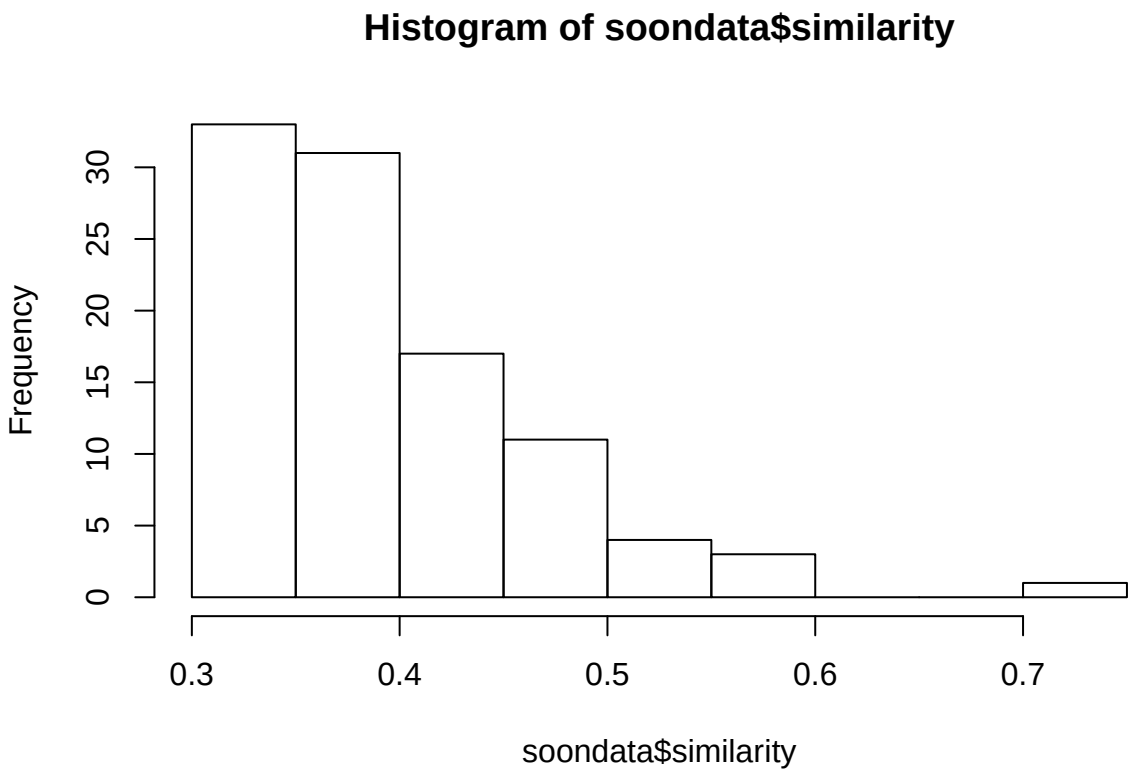
Below: A histogram of word similarities to “long”.



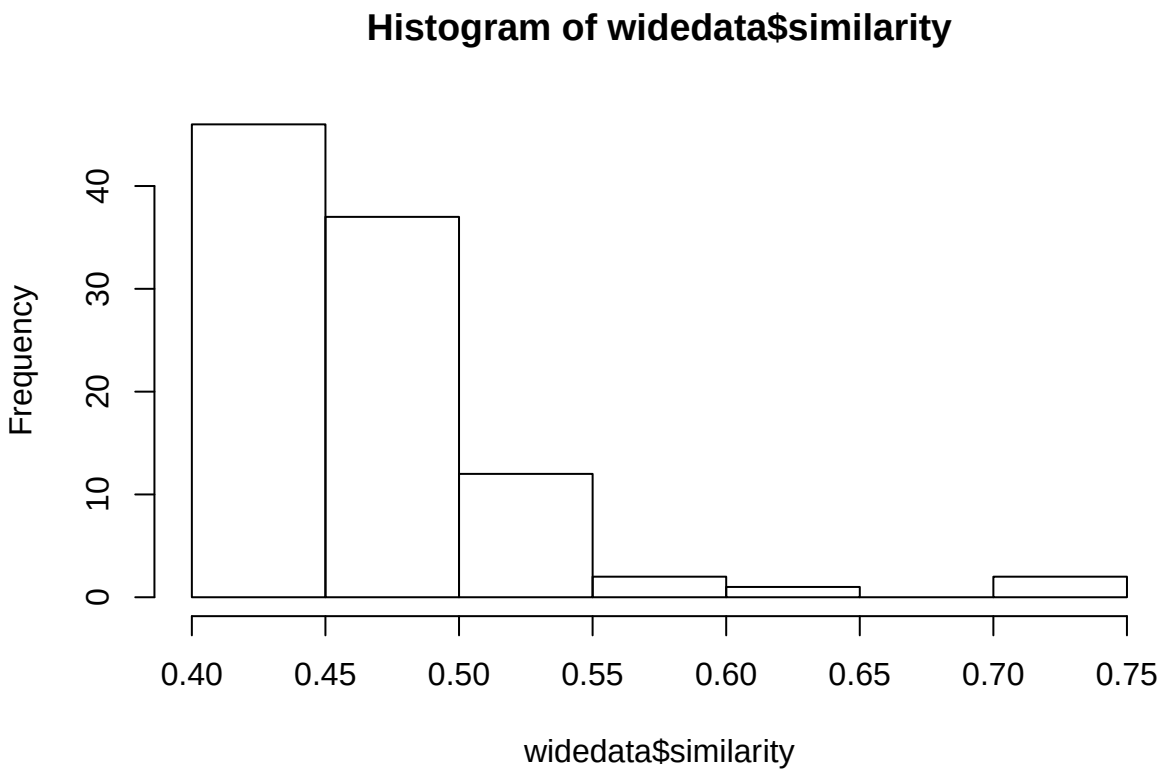
Below: A histogram of word similarities to “short”.



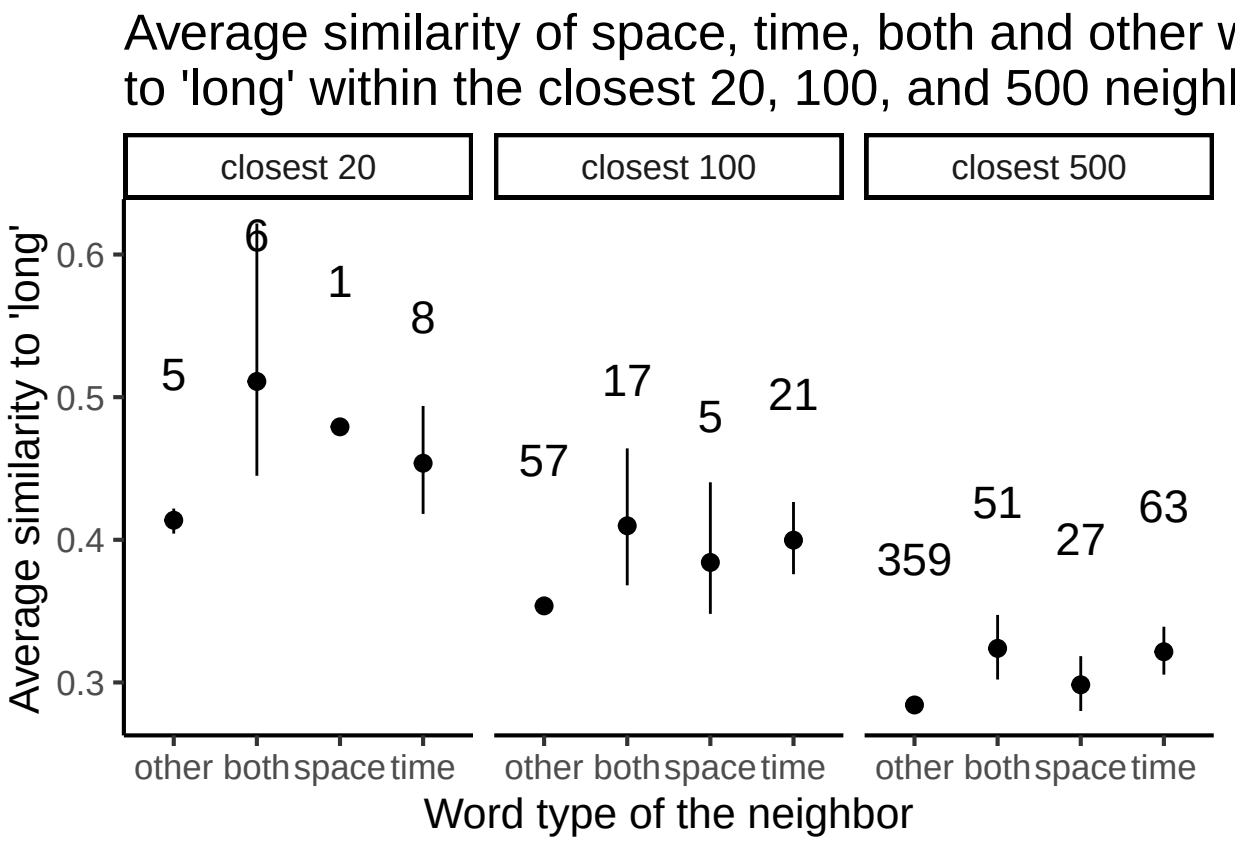
Below: A histogram of word similarities to “soon”.



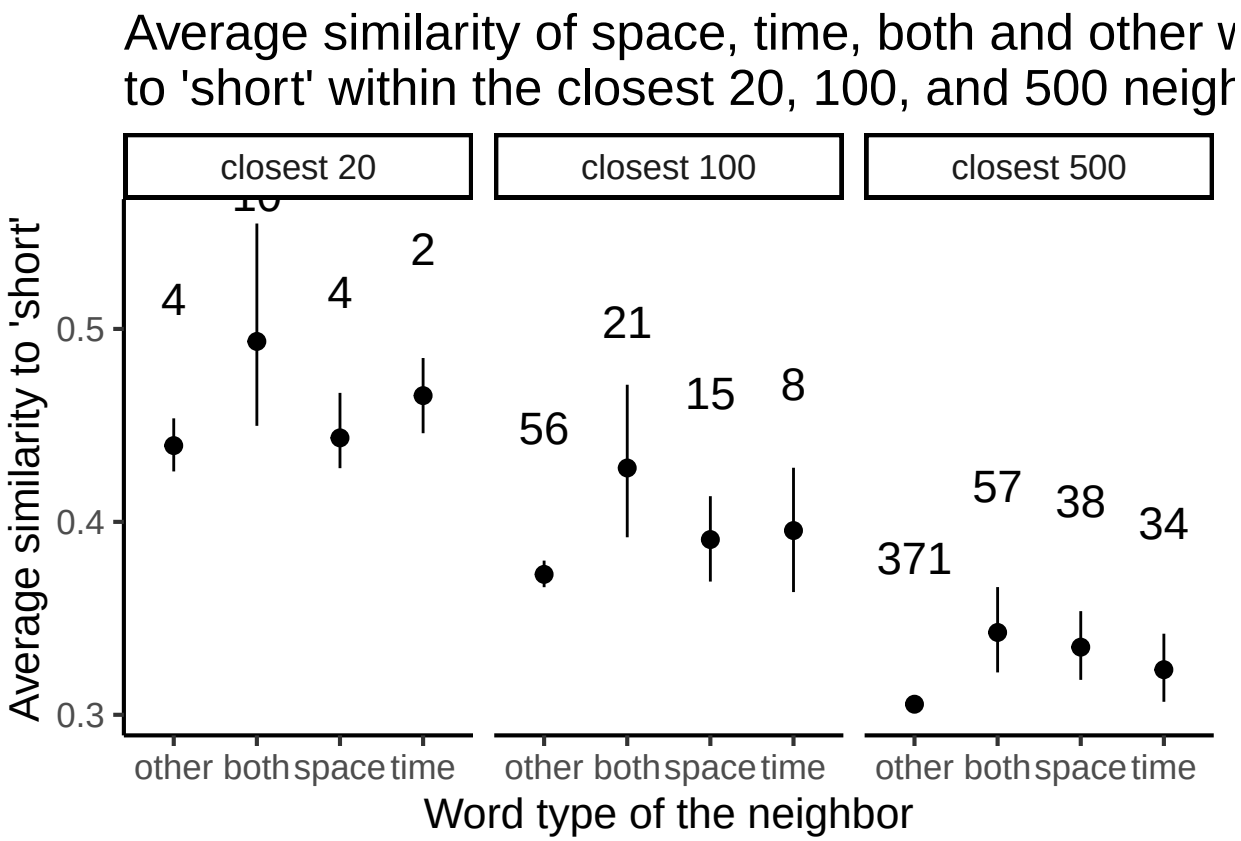
Below: A histogram of word similarities to “wide”.



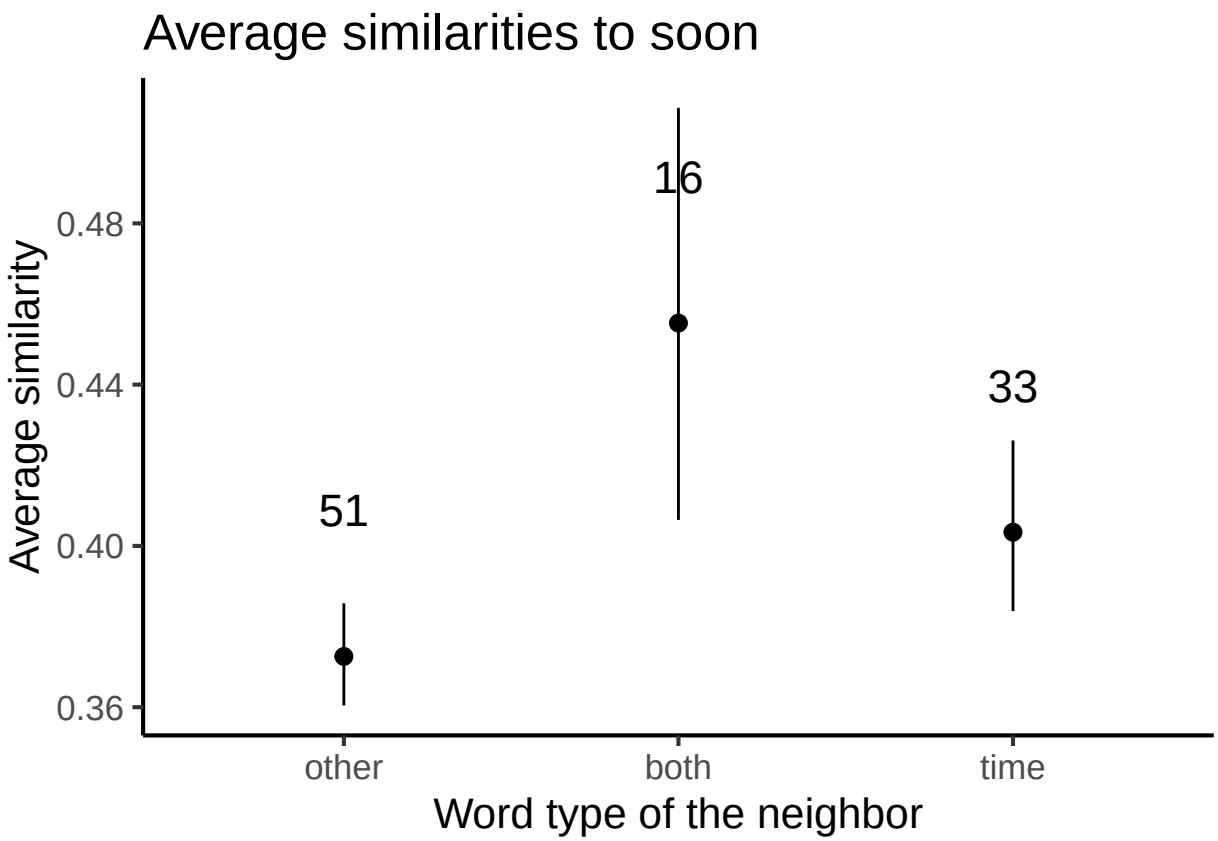
Below: A plot of mean similarities of space, time, both and other words to "long" within the closest 20, 100, and 500 neighbors. The numbers on the plot denote the number of words of that type in that range of neighbors.



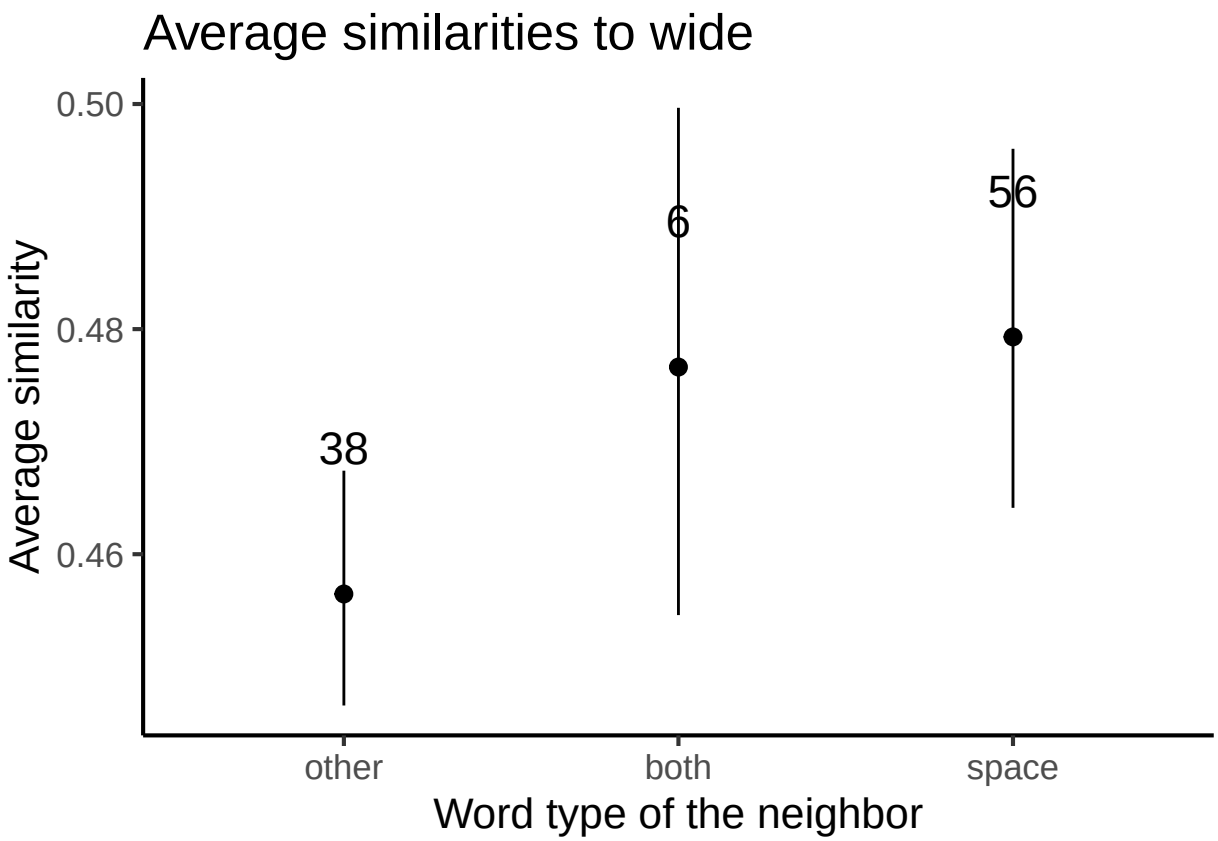
Below: A plot of mean similarities of space, time, both and other words to "short" within the closest 20, 100, and 500 neighbors.



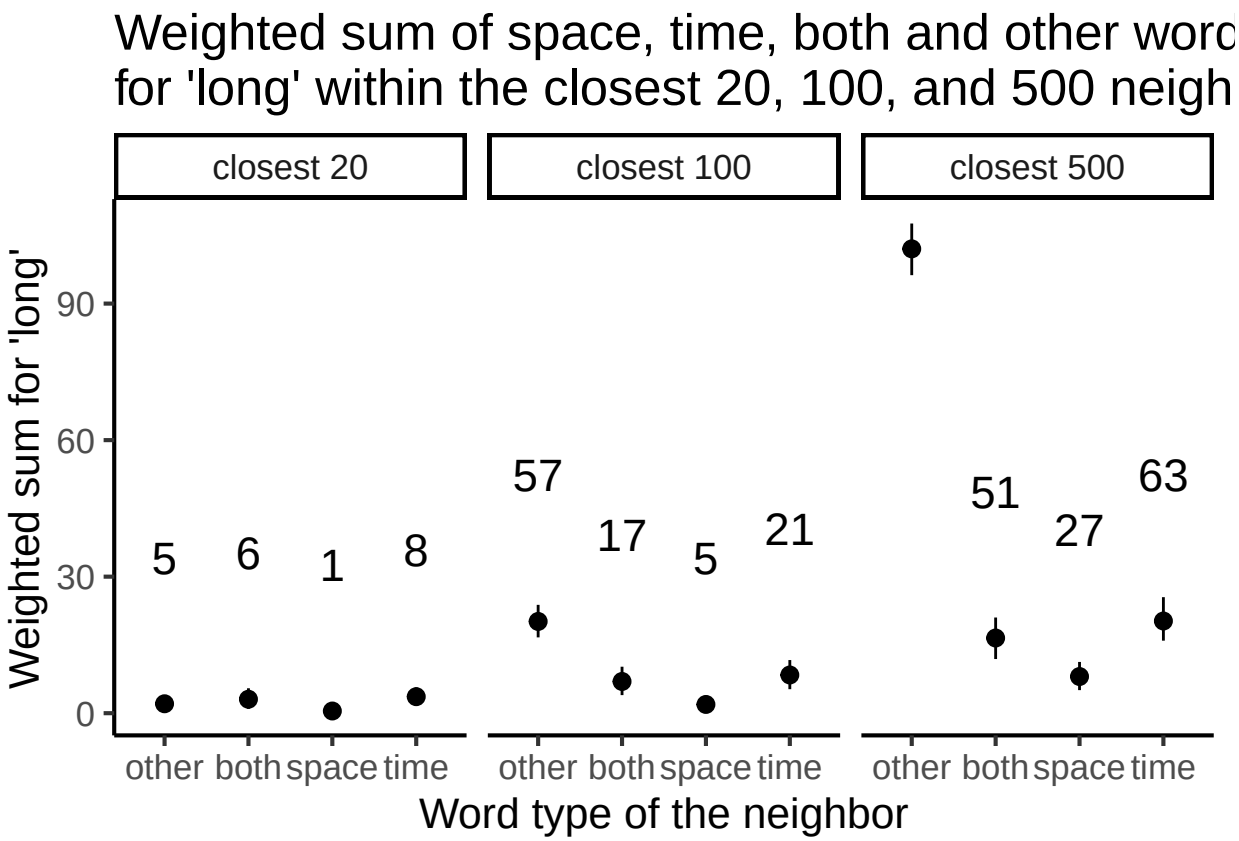
Below: Mean similarities of space, time, both and other words to "soon". Note: there were no pure space words for 'soon'.



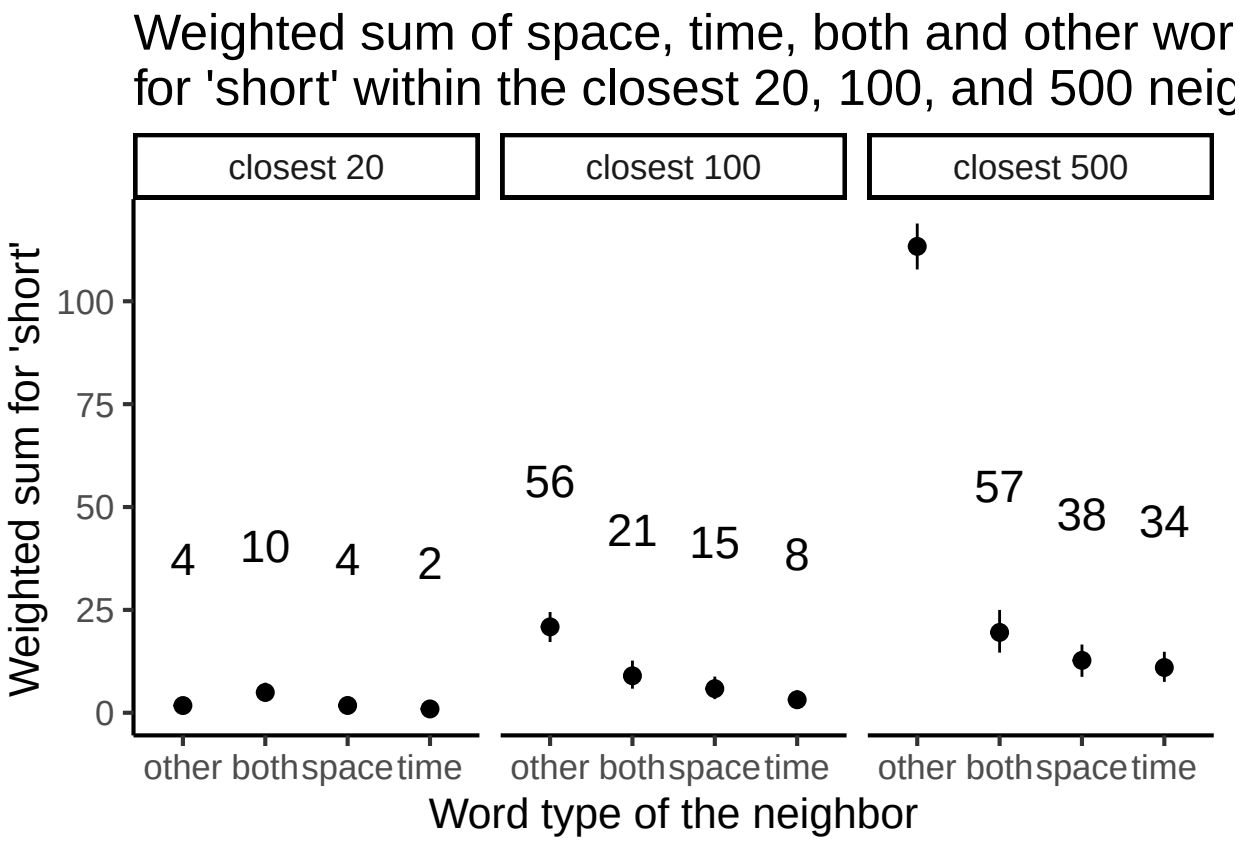
Below: Mean similarities of space, time, both and other words to “wide”. Note: there were no pure time words for ‘wide’.



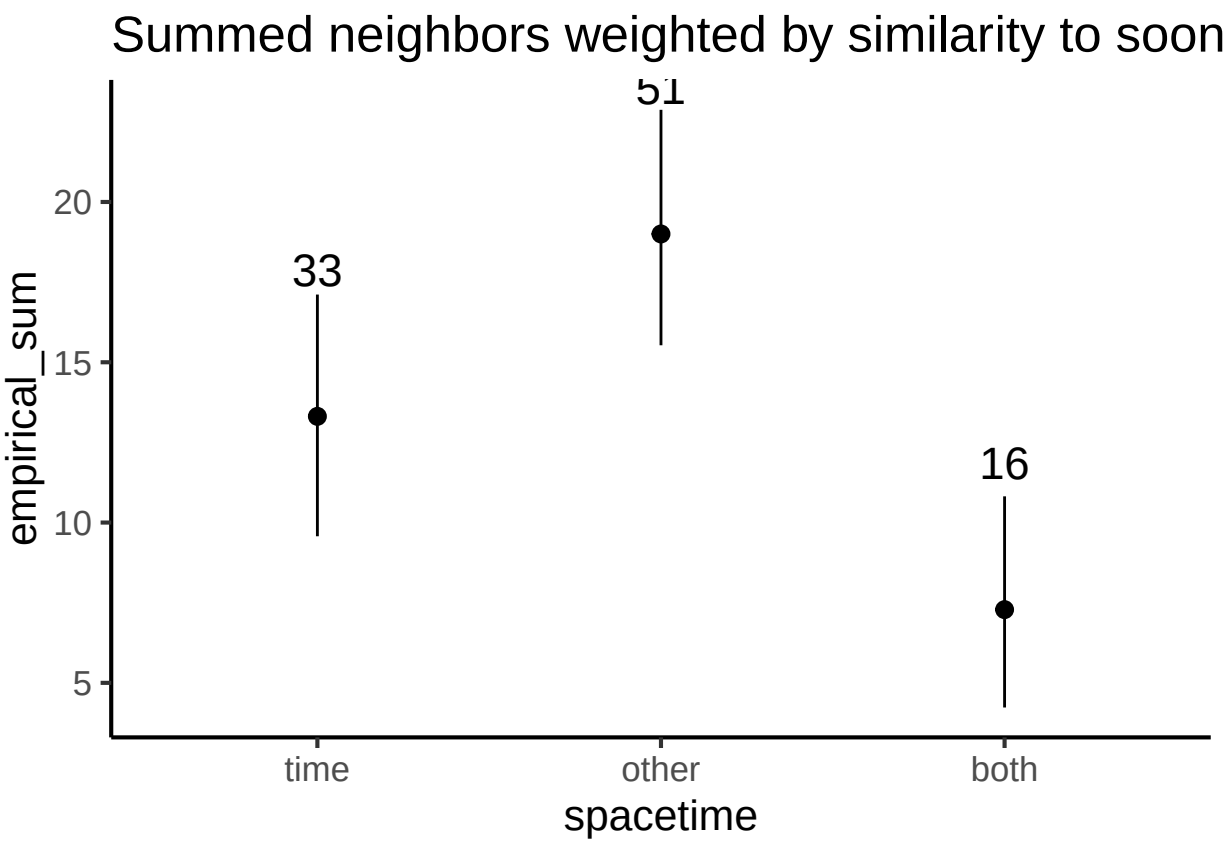
Below: A plot of weighted sums of space, time, both and other words for "long" within the closest 20, 100, and 500 neighbors. The numbers on the plot denote the number of words of that type in that range of neighbors.



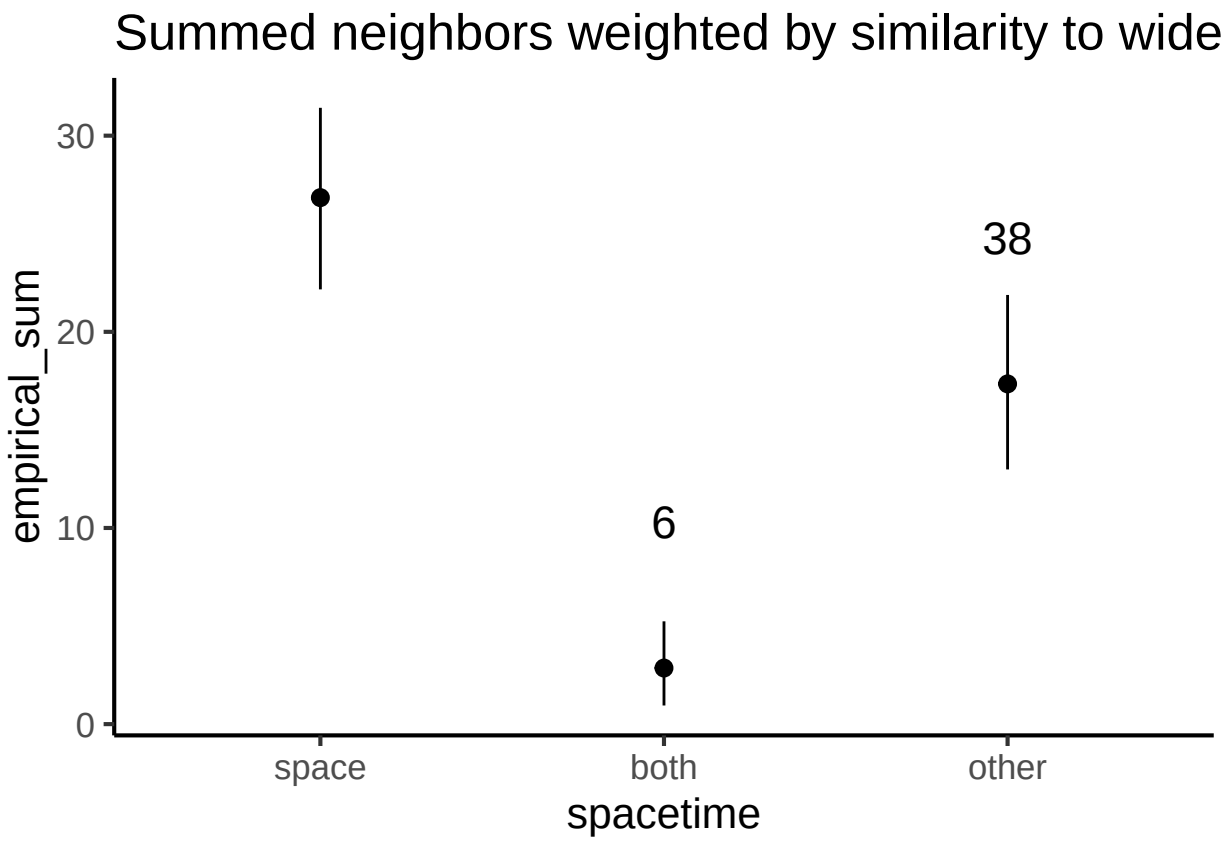
Below: A plot of weighted sums of space, time, both and other words for "short" within the closest 20, 100, and 500 neighbors. The numbers on the plot denote the number of words of that type in that range of neighbors.



Below: Summed similarities of space, time, both and other words to “soon”. Note: there were no pure space words for ‘soon’.



Below: Summed similarities of space, time, both and other words to “wide”. Note: there were no pure time words for ‘wide’.



Below: Linear model predicting similarity to "long" by word type (space, time, etc.) in the most similar 100 words, where we're more likely to see differences. Space is the reference category.

```
##
## Call:
## glm(formula = similarity ~ spacetime, family = gaussian, data = longmodeldata)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.08337  -0.02370  -0.00912   0.01346   0.33552
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.38415    0.02433  15.792   <2e-16 ***
## spacetimeother -0.03047    0.02537  -1.201    0.233
## spacetimetime   0.01561    0.02707   0.577    0.566
## spacetimeboth   0.02569    0.02767   0.928    0.356
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.002958698)
##
##     Null deviance: 0.34383  on 99  degrees of freedom
## Residual deviance: 0.28403  on 96  degrees of freedom
## AIC: -292.6
##
```

```
## Number of Fisher Scoring iterations: 2
```

Below: Linear model predicting similarity to "short" by word type (space, time, etc.) in the most similar 100 words, where we're more likely to see differences. Space is the reference category.

```
##
## Call:
## glm(formula = similarity ~ spacetime, family = gaussian, data = shortmodeldata)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.07889  -0.02731  -0.00803   0.02449   0.31743
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.390836   0.013160  29.698   <2e-16 ***
## spacetimeother -0.018057   0.014819  -1.219   0.2260
## spacetimeboth   0.037086   0.017231   2.152   0.0339 *
## spacetimetime   0.004661   0.022315   0.209   0.8350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.002597968)
##
##     Null deviance: 0.29634  on 99  degrees of freedom
## Residual deviance: 0.24940  on 96  degrees of freedom
## AIC: -305.6
##
## Number of Fisher Scoring iterations: 2
```

Functions to get the counts

Below: Empirical values for the number of space words, the number of time words and the difference between the two within the first 100 nearest neigbors of long.

Below: Empirical values for the number of space words, the number of time words and the difference between the two within the first 100 nearest neigbors of short

Below: Bootstrapping (n=100) samples of number of space words, time words, and their difference from the sample of 6000 random words.

Import the random sample from the corpus with 6000 words.

Generate the null distribution for number of space words, time words, and their difference.

Below: Same analyses with the 500 observed sample and bootstrapped samples with N=500

Below: Empirical values for the number of space words, the number of time words and the difference between the two within the first 500 nearest neigbors of long.

Below: Empirical values for the number of space words, the number of time words and the difference between the two within the first 500 nearest neigbors of short

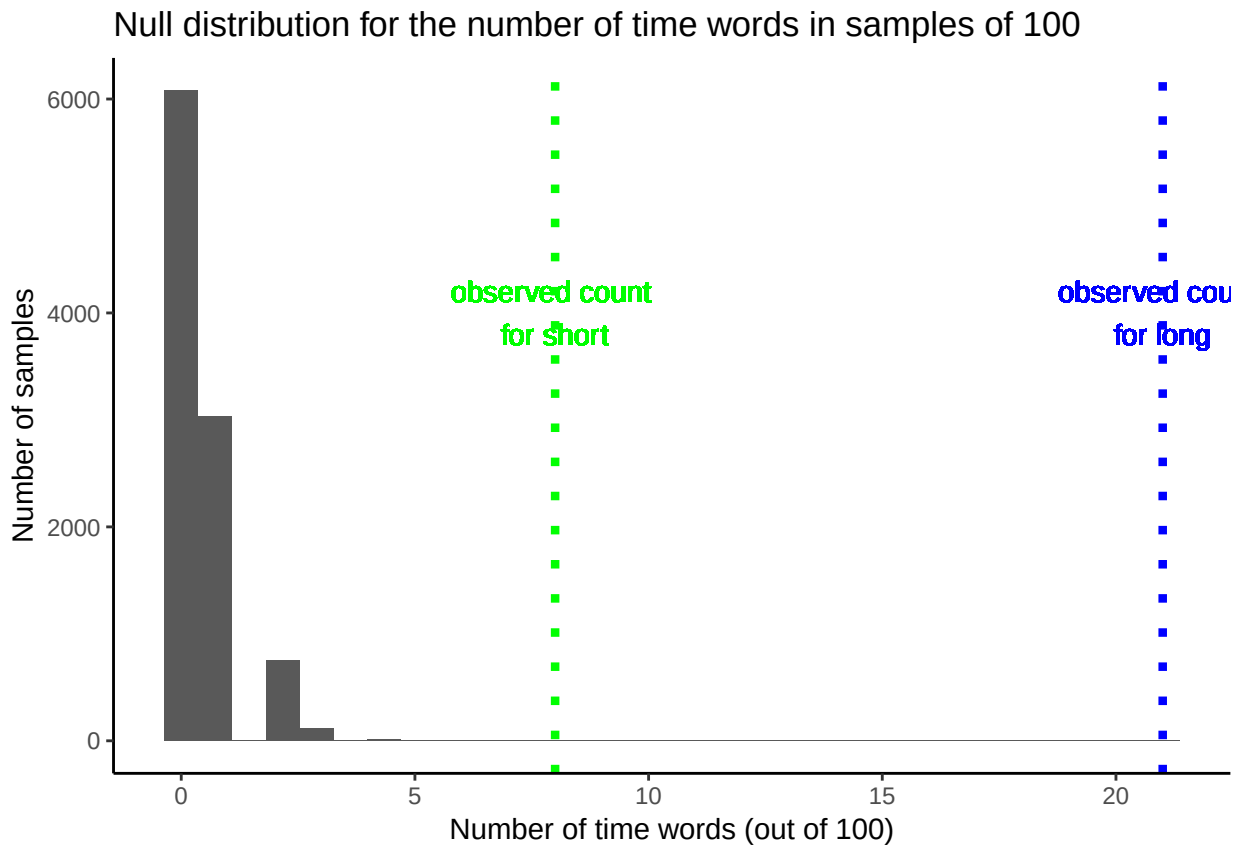
Generate the null distribution for number of space words, time words, and their difference – samples with n = 500

TIME WORDS: among samples of 100 and 500.

```
## Warning: Ignoring unknown parameters: text
```

```
## Warning: Ignoring unknown parameters: text
```

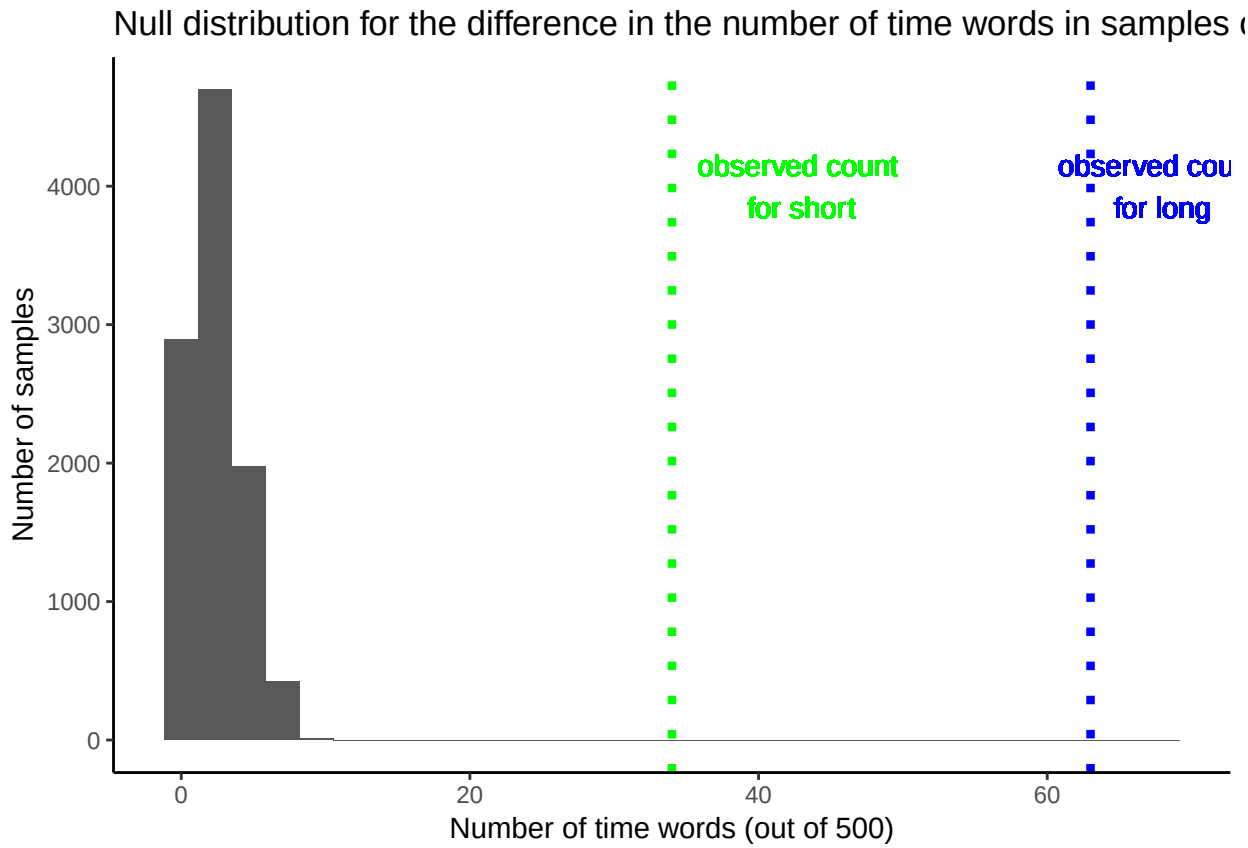
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## Warning: Ignoring unknown parameters: text
```

```
## Warning: Ignoring unknown parameters: text
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

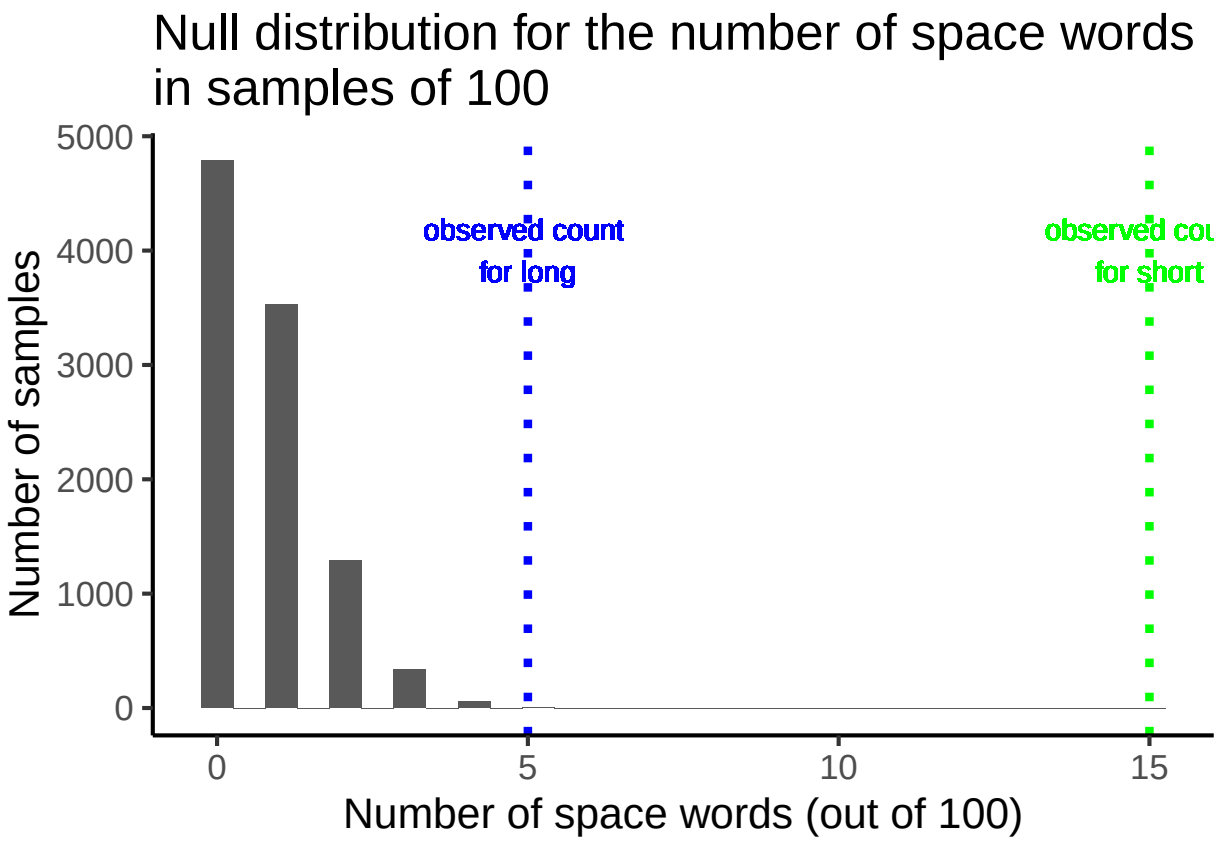


SPACE WORDS: among samples of 100 and 500.

```
## Warning: Ignoring unknown parameters: text
```

```
## Warning: Ignoring unknown parameters: text
```

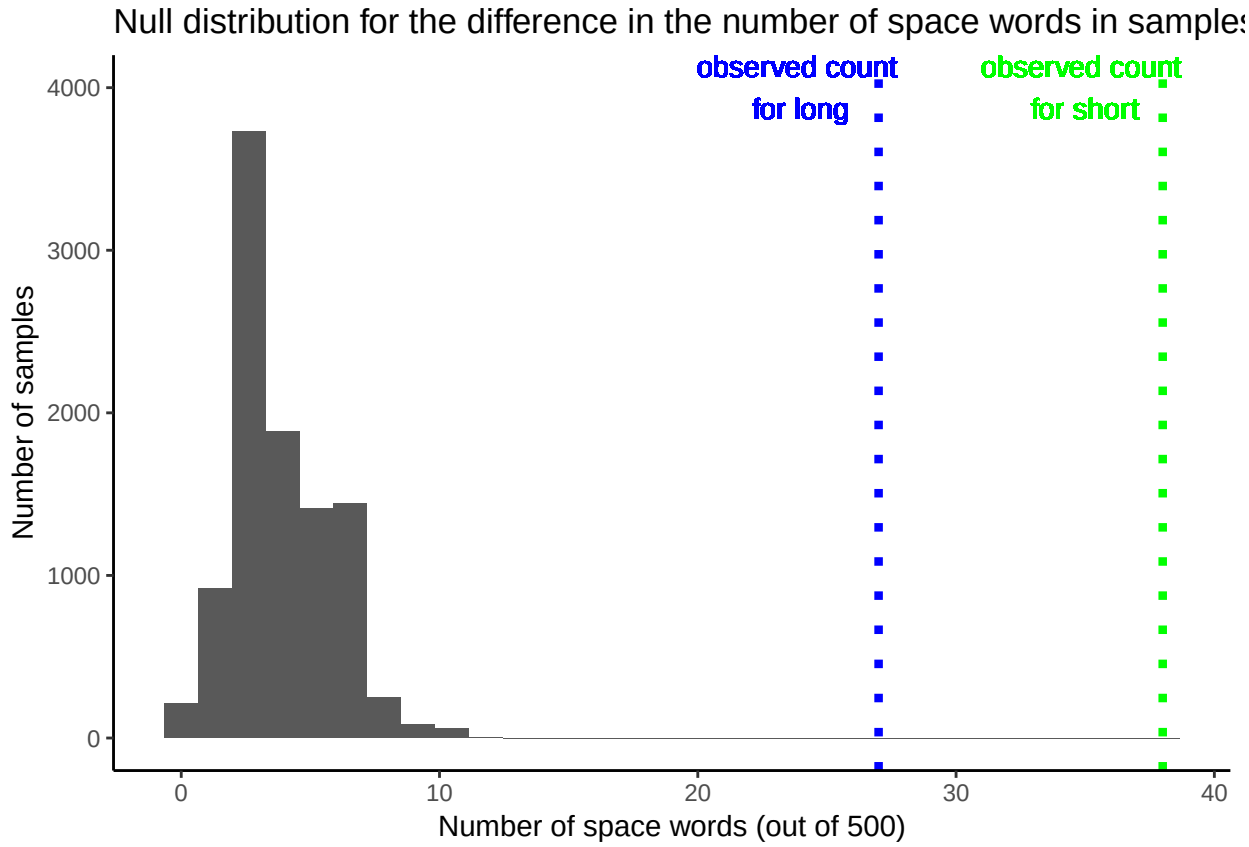
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## Warning: Ignoring unknown parameters: text

## Warning: Ignoring unknown parameters: text

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

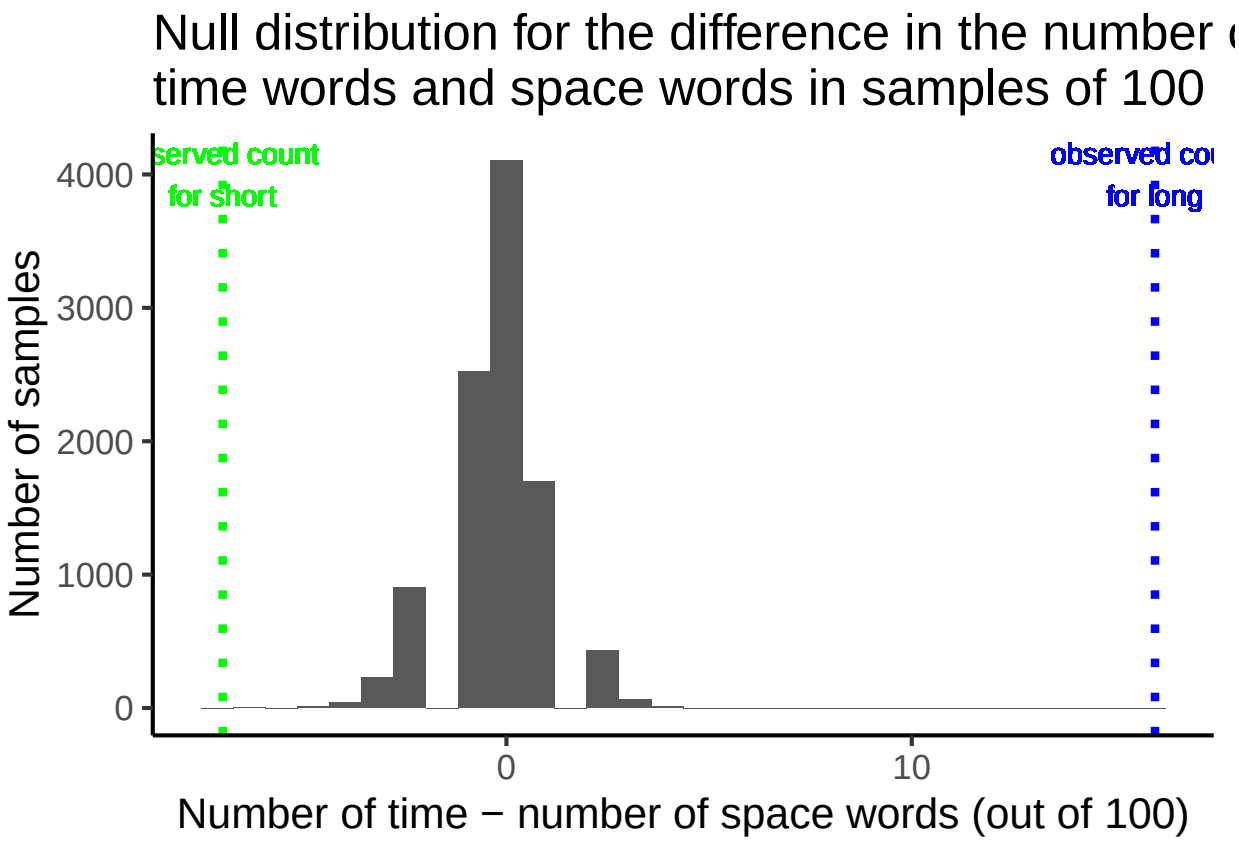


TIME WORDS - SPACE WORDS: among samples of 100 and 500.

```
## Warning: Ignoring unknown parameters: text
```

```
## Warning: Ignoring unknown parameters: text
```

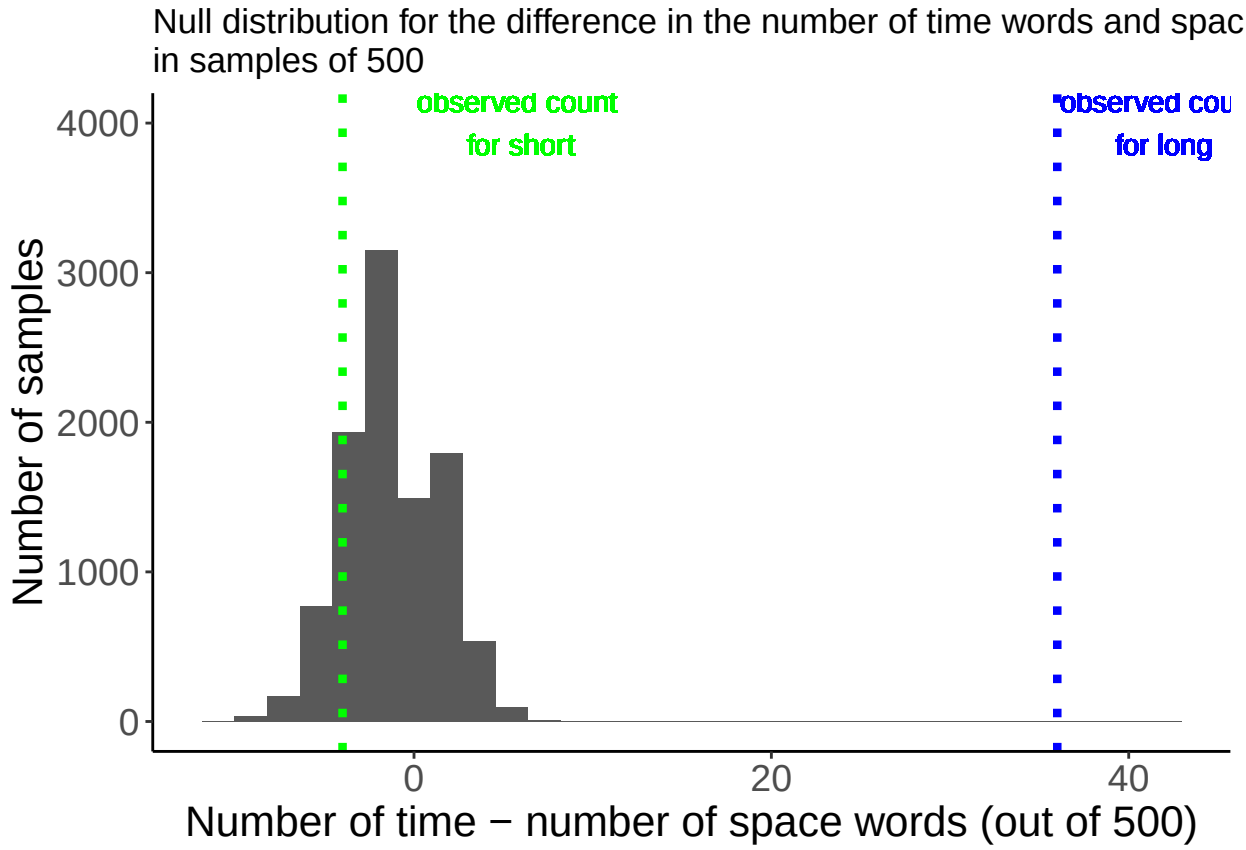
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## Warning: Ignoring unknown parameters: text

## Warning: Ignoring unknown parameters: text

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Stats with samples N=500

For long

```
## [1] 0
```

```
## [1] 0
```

```
## [1] 0
```

For short

```
## [1] 0.9026
```

```
## [1] 0
```

```
## [1] 0
```