

spacetimeadultanalysis

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

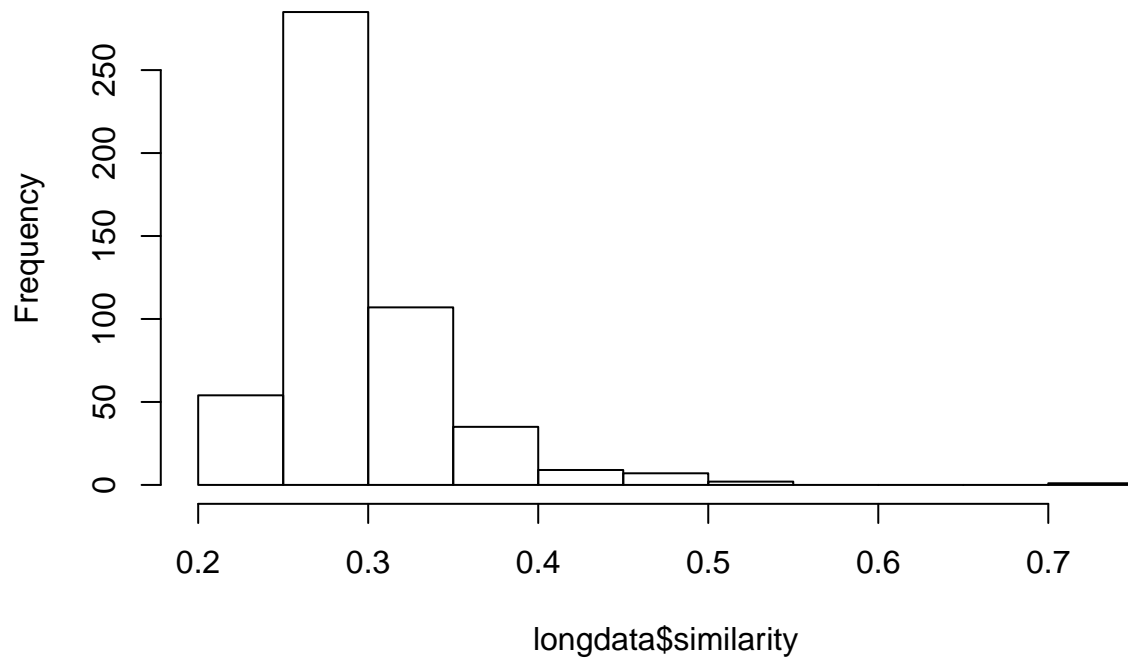
```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

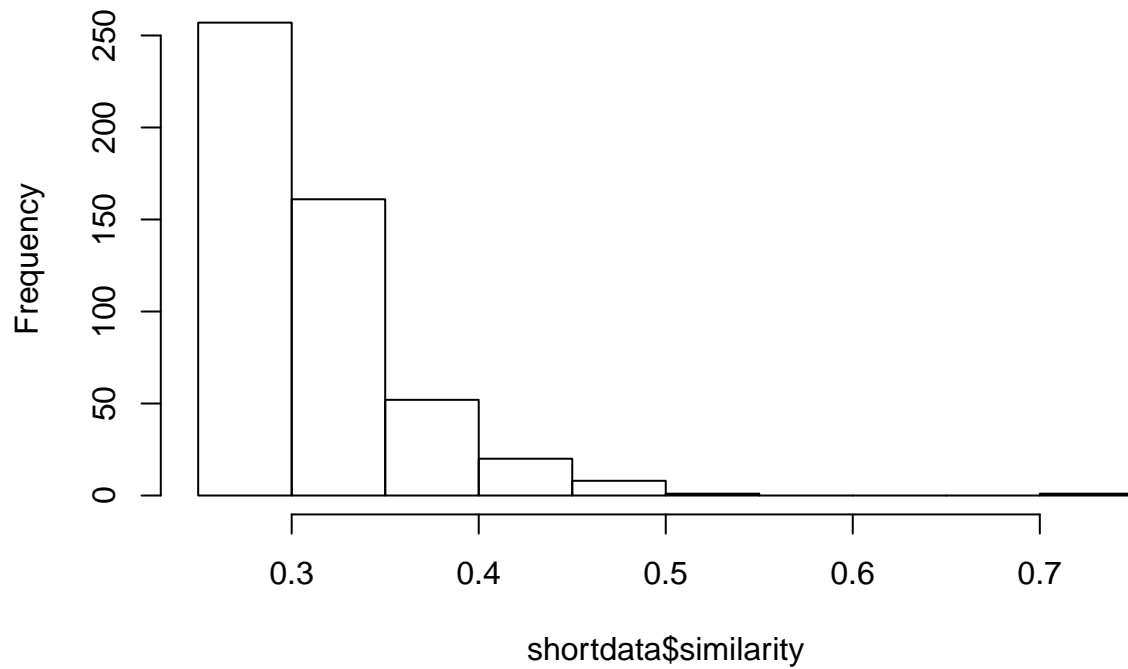
Below: A histogram of word similarities to “long”.

Histogram of longdata\$similarity



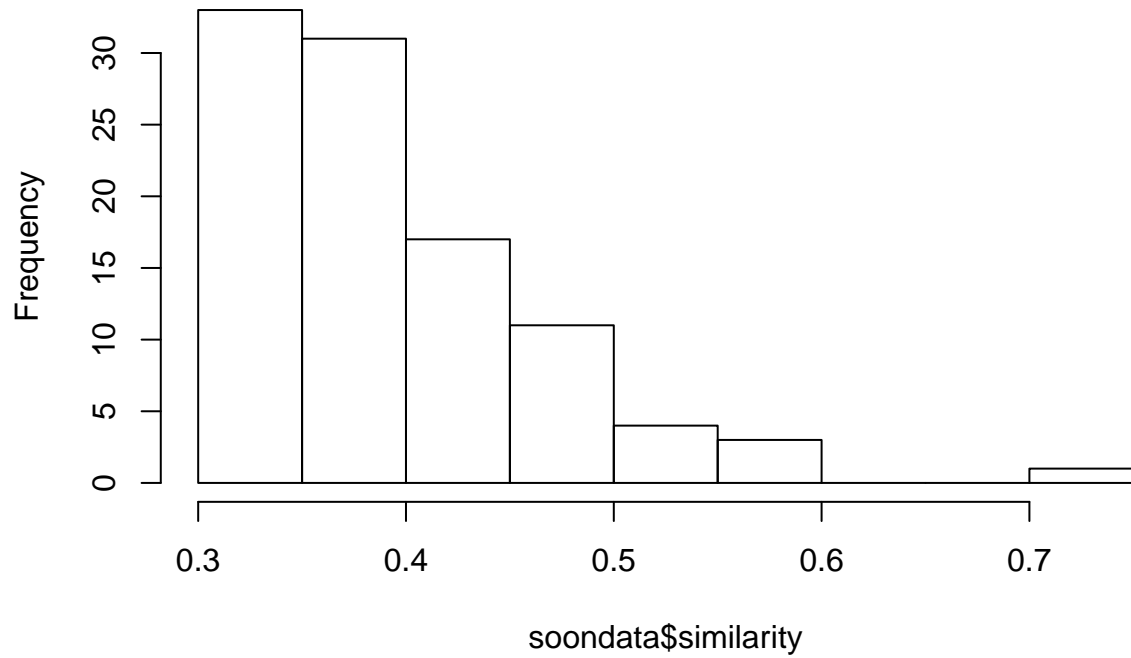
Below: A histogram of word similarities to “short”.

Histogram of shortdata\$similarity



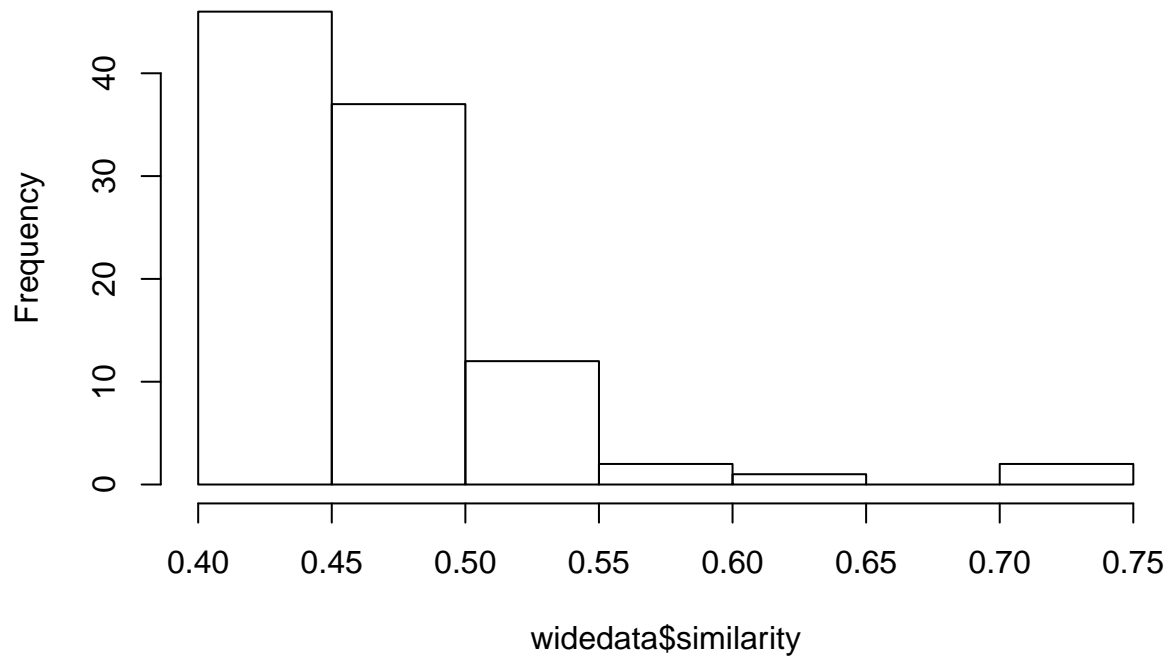
Below: A histogram of word similarities to “soon”.

Histogram of soondata\$similarity

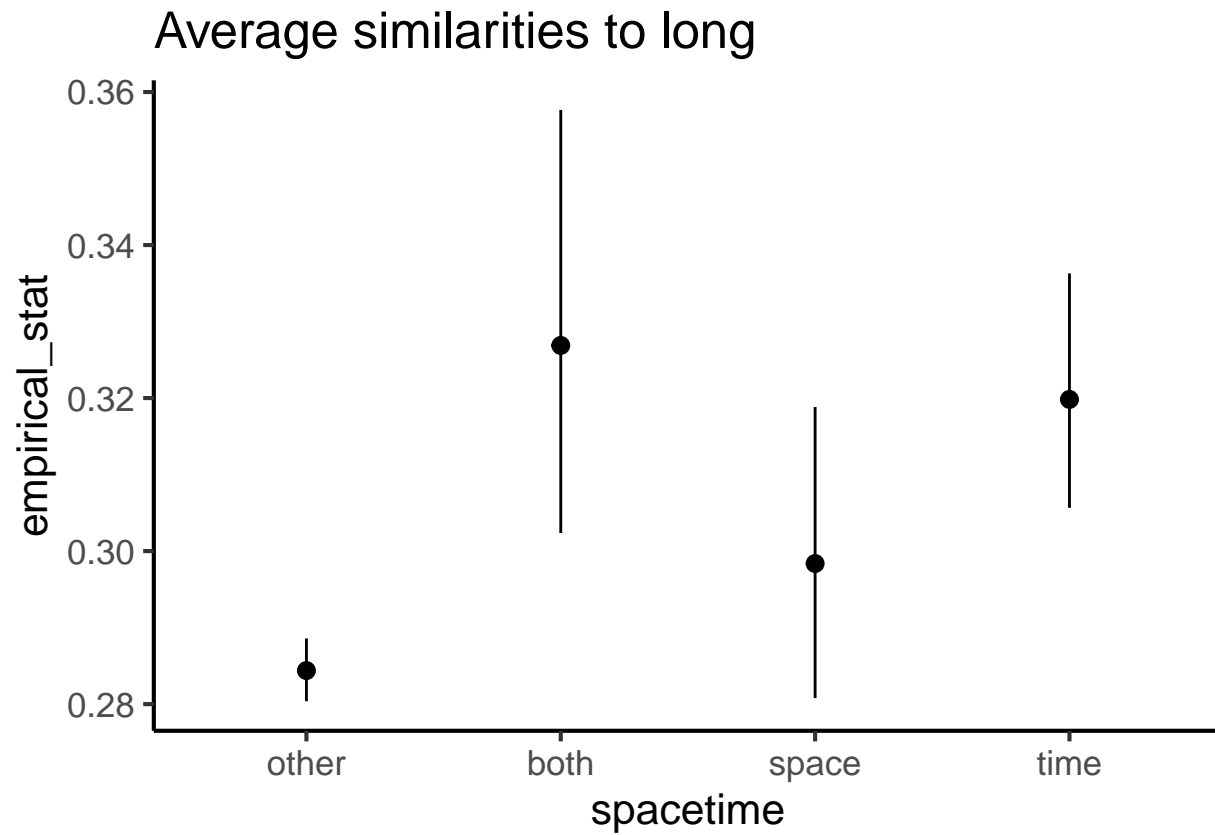


Below: A histogram of word similarities to “wide”.

Histogram of widedata\$similarity

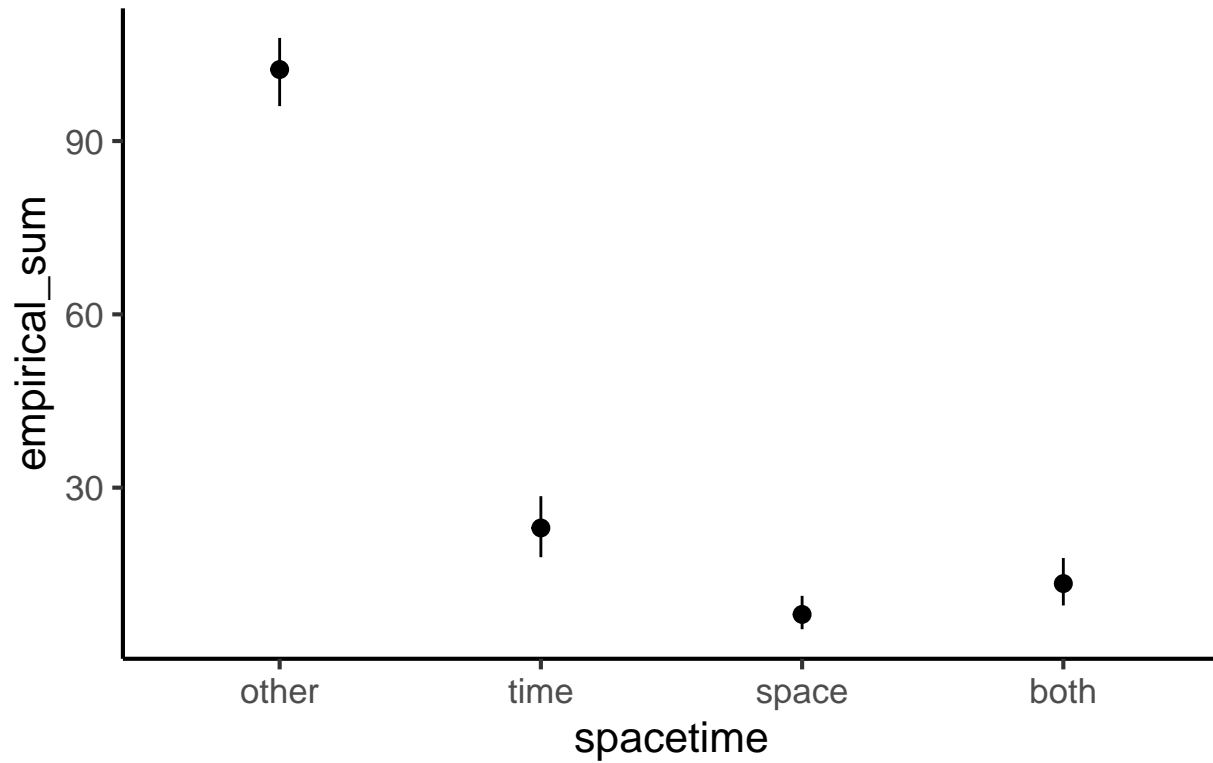


Below: Mean similarities of space, time, both and other words to “long”. Error bars here and throughout are 95% bootstrapped confidence intervals.



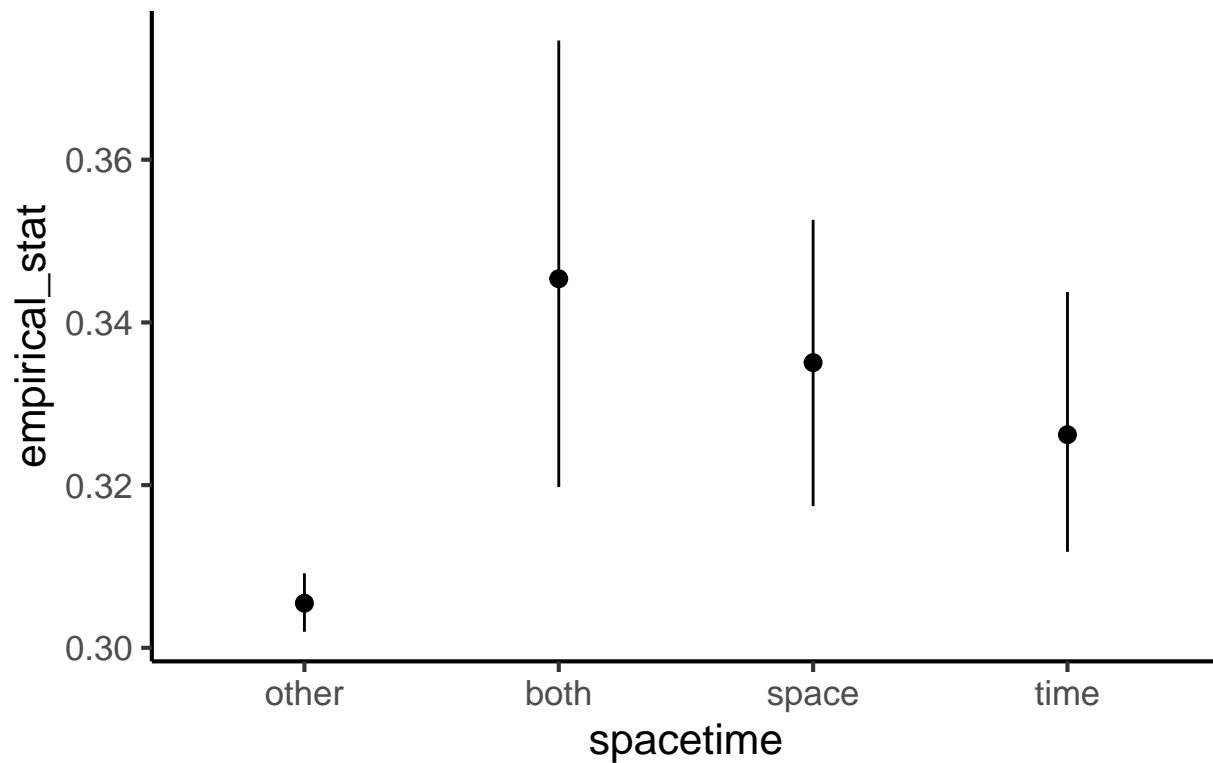
Below: Summed similarities of space, time, both and other words to “long”. Error bars here and throughout are 95% bootstrapped confidence intervals.

Summed neighbors weighted by similarity to long

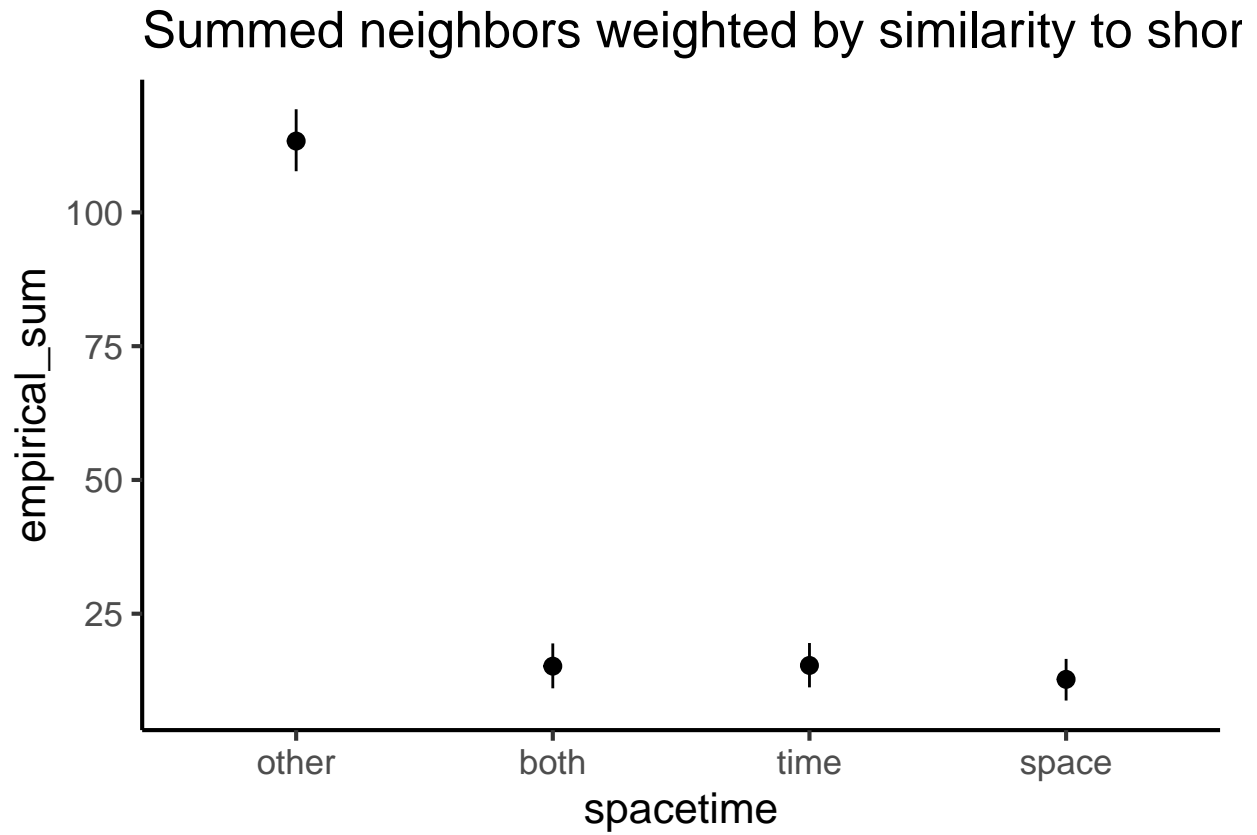


Below: Mean similarities of space, time, both and other words to “short”.

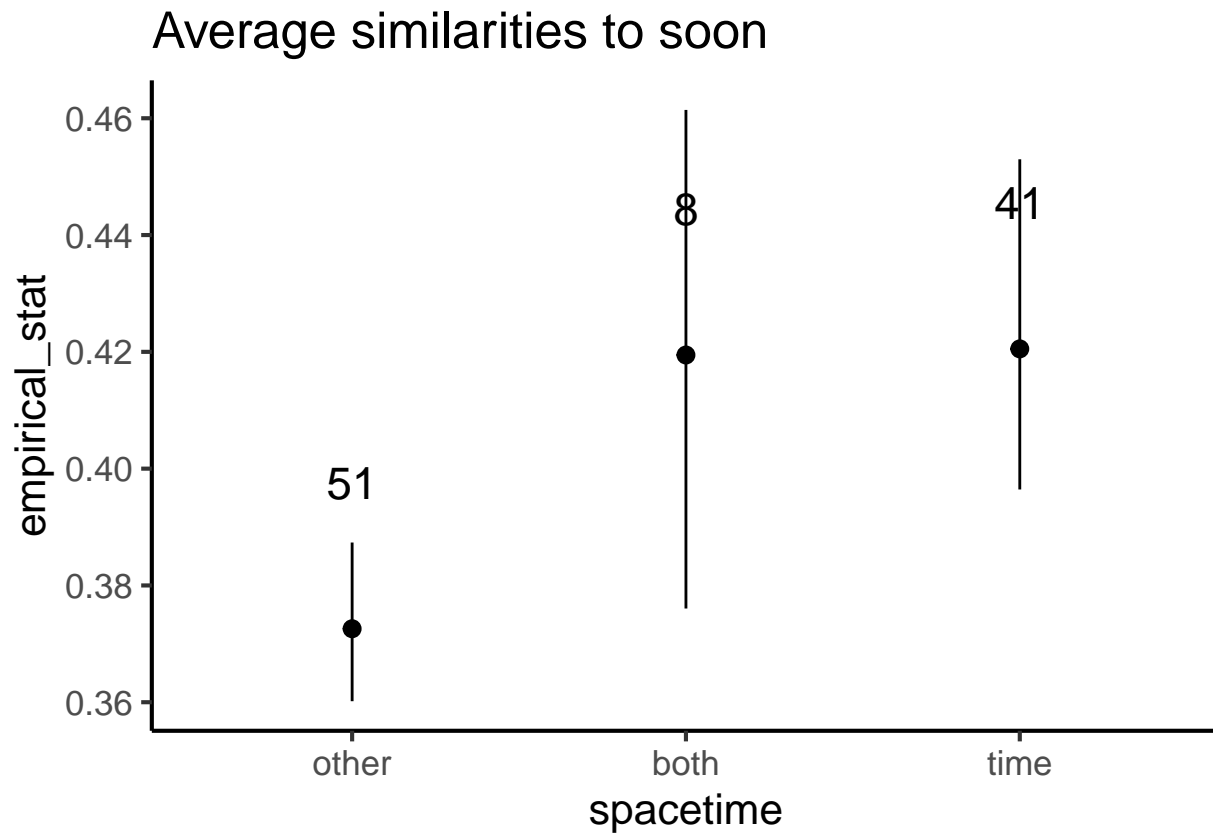
Average similarities to short



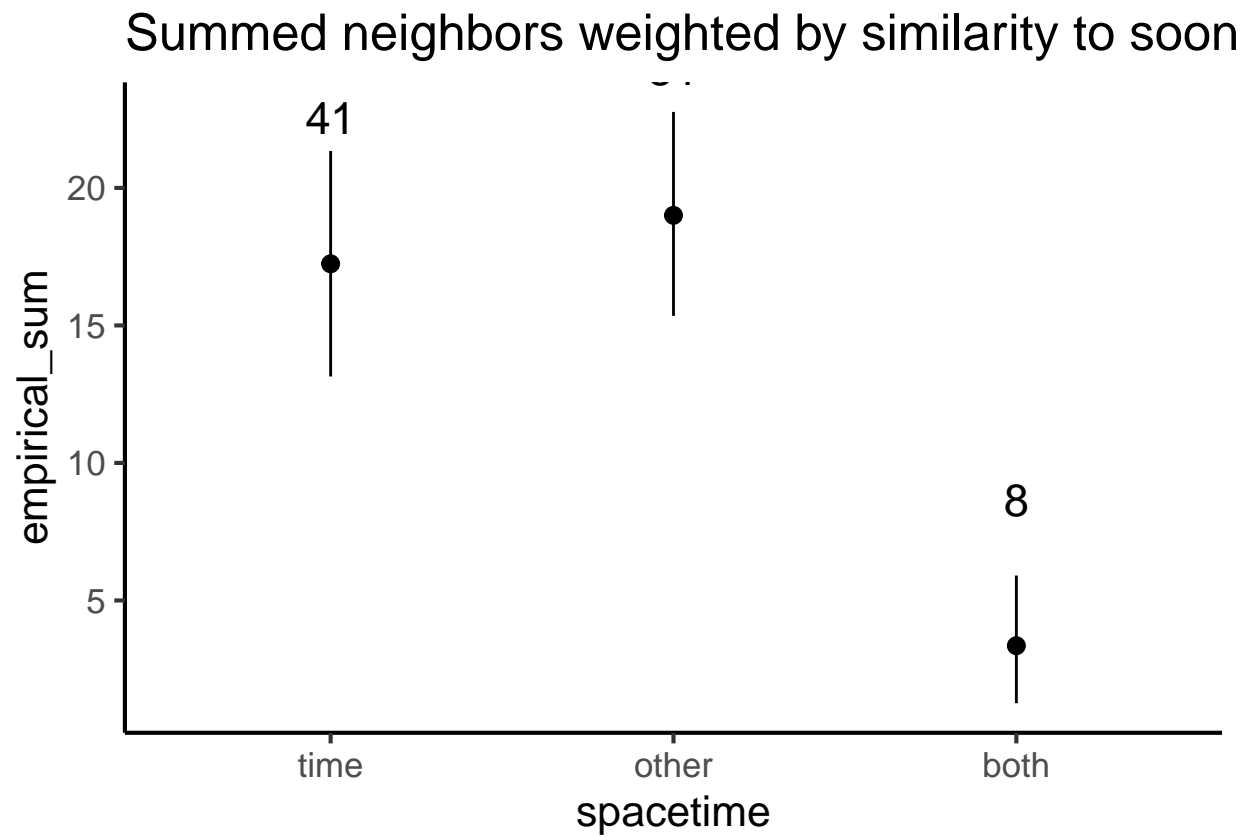
Below: Summed similarities of space, time, both and other words to “short”.



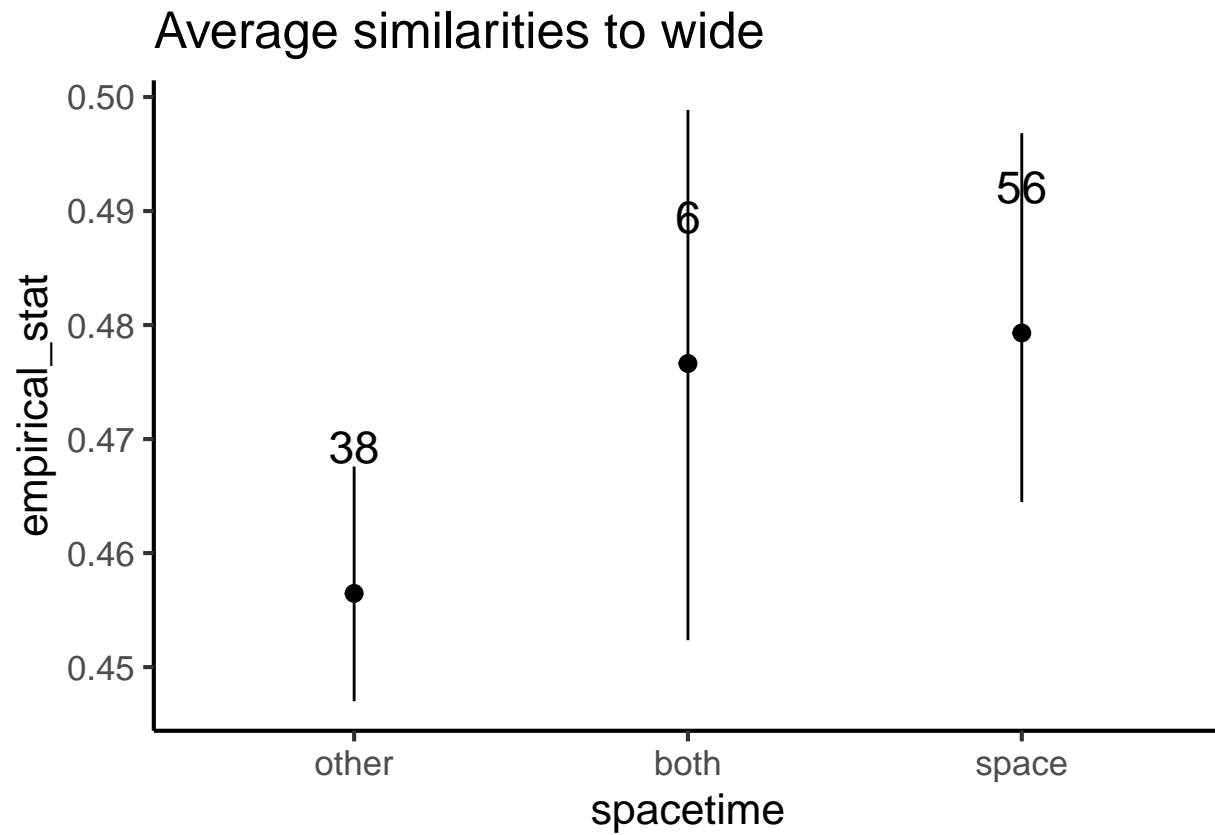
Below: Mean similarities of space, time, both and other words to “soon”. Note: there were no pure space words for ‘soon’.



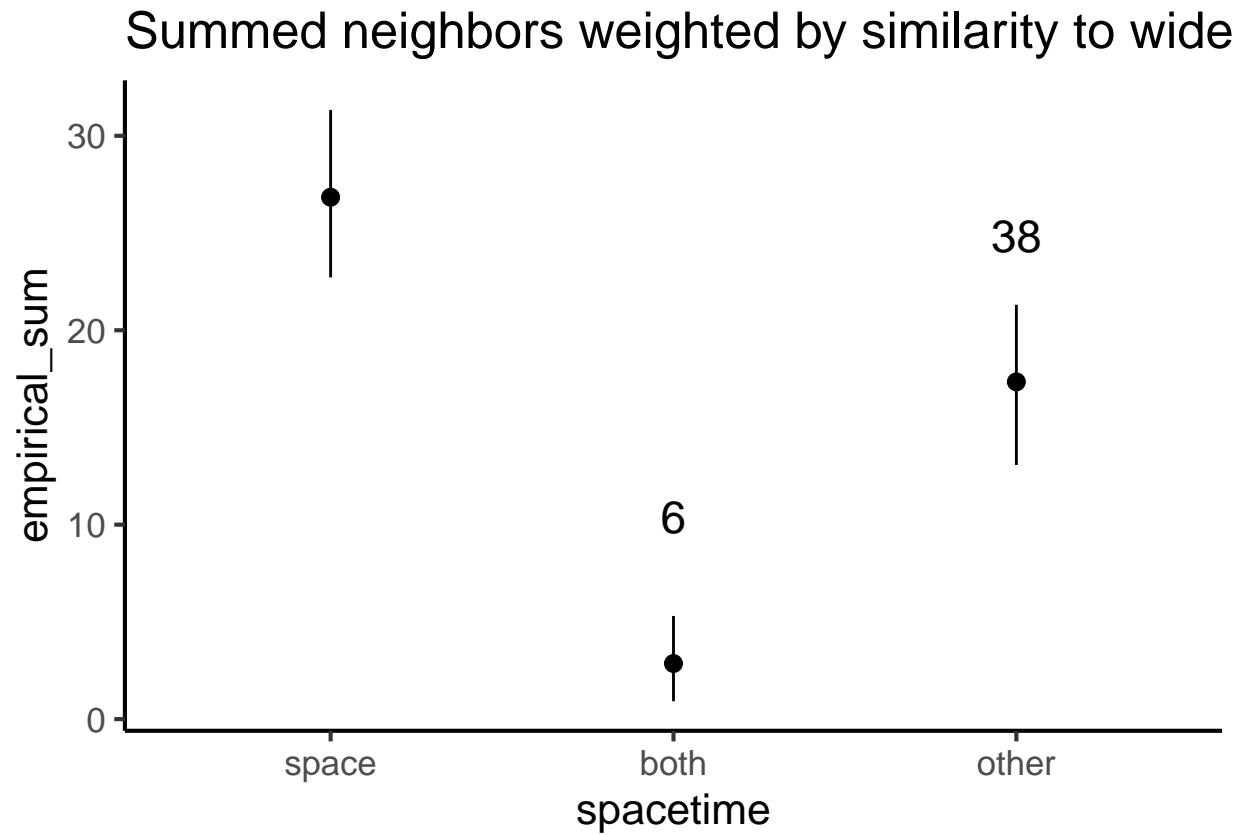
Below: Summed similarities of space, time, both and other words to “soon”. Note: there were no pure space words for ‘soon’.



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



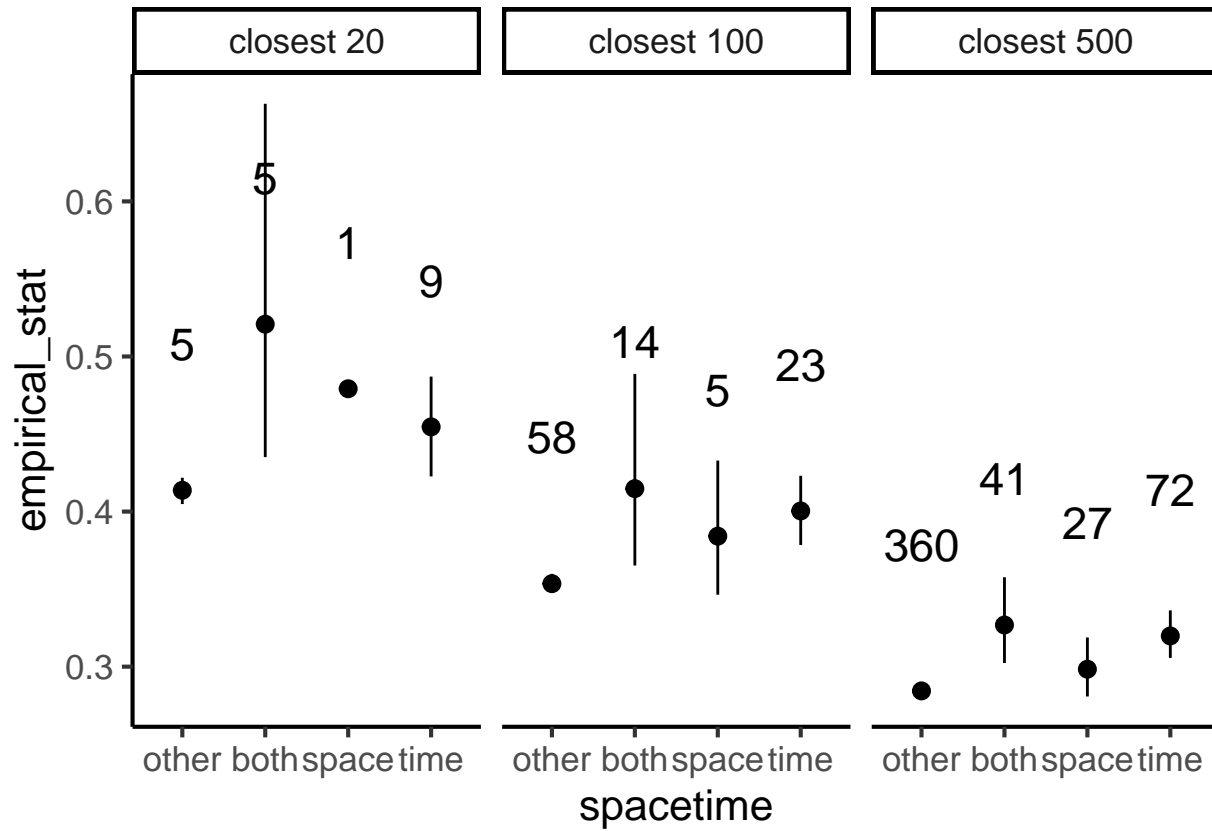
Below: Summed similarities of space, time, both and other words to “wide”. Note: there were no pure time words for ‘wide’.



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.

```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```

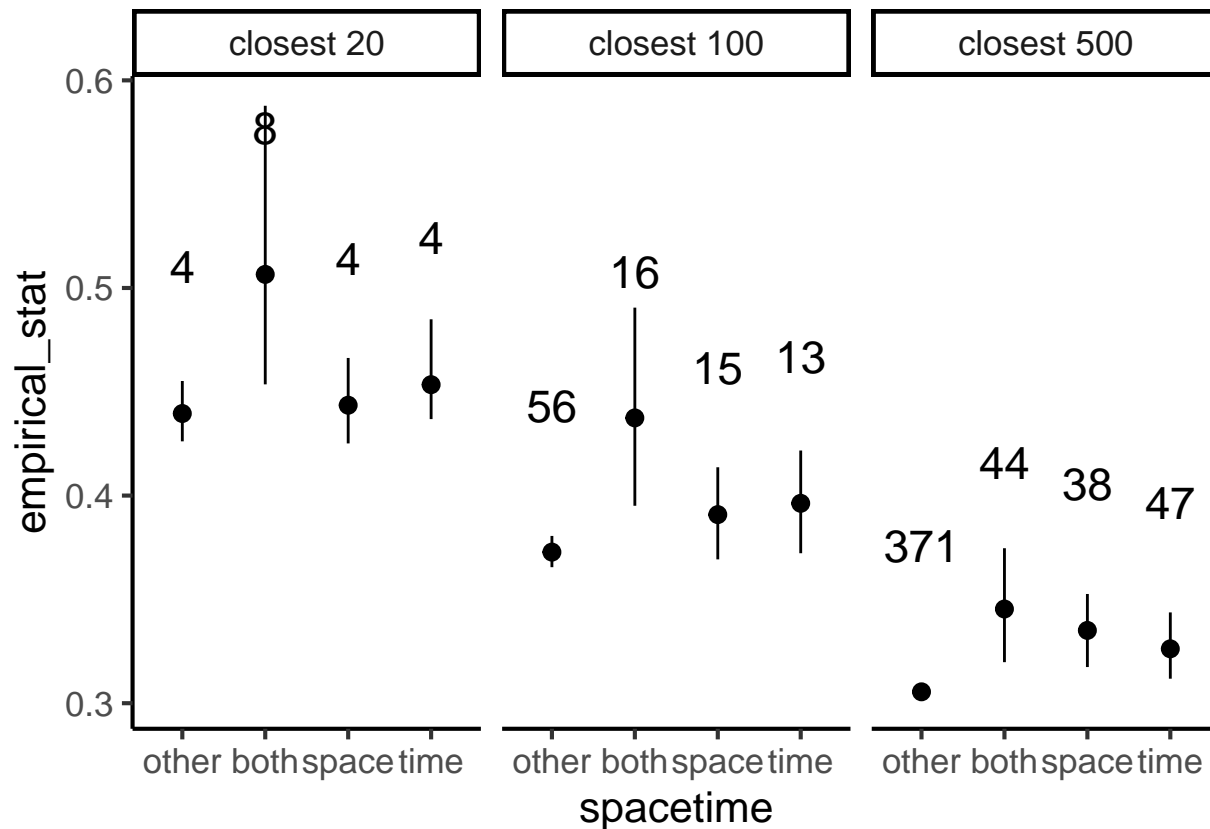
```
## Warning: `cols` is now required.  
## Please use `cols = c(strap)`
```



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

```
## Warning: `cols` is now required.
## Please use `cols = c(strap)`
```

```
## Warning: `cols` is now required.
## Please use `cols = c(strap)`
```



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
## -0.08834 -0.02335 -0.00867  0.01301  0.33055
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.38415    0.02414  15.914 <2e-16 ***
## spacetimeother -0.03061    0.02516  -1.217   0.227
## spacetime     0.01626    0.02663   0.610   0.543
## spacetimeboth  0.03066    0.02812   1.090   0.278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.002913571)
##
##      Null deviance: 0.34383  on 99  degrees of freedom
## Residual deviance: 0.27970  on 96  degrees of freedom
## AIC: -294.13
##
## 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
## -0.082279  -0.027310  -0.008033   0.025199   0.307933
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.390836   0.013000  30.065  <2e-16 ***
## spacetimeother -0.018057   0.014638  -1.234   0.2204
## spacetimeboth   0.046585   0.018095   2.574   0.0116 *
## spacetime      0.005441   0.019079   0.285   0.7761
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.002534941)
##
##      Null deviance: 0.29634  on 99  degrees of freedom
## Residual deviance: 0.24335  on 96  degrees of freedom
## AIC: -308.05
##
## 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 neighbors 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 neighbors of short

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

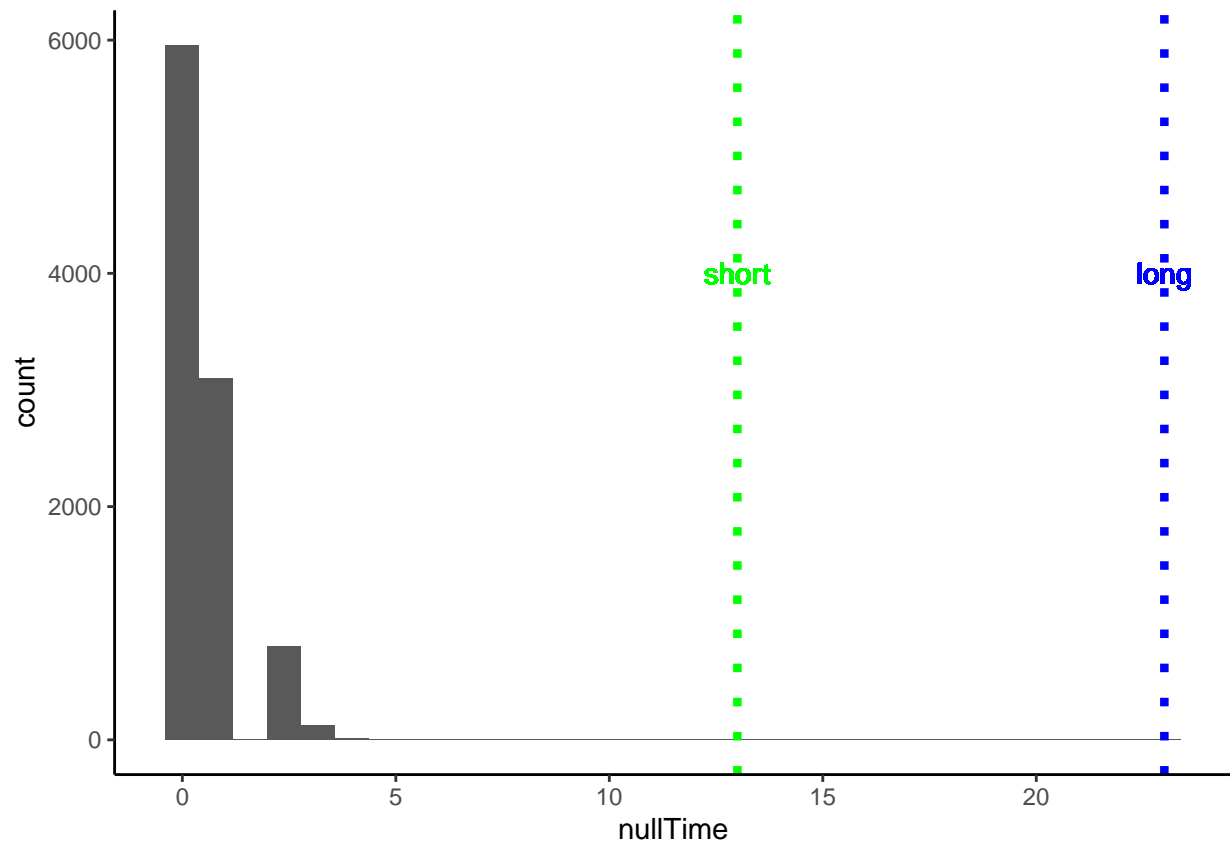
Import the random sample from the corpus with 6000 words.

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

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

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

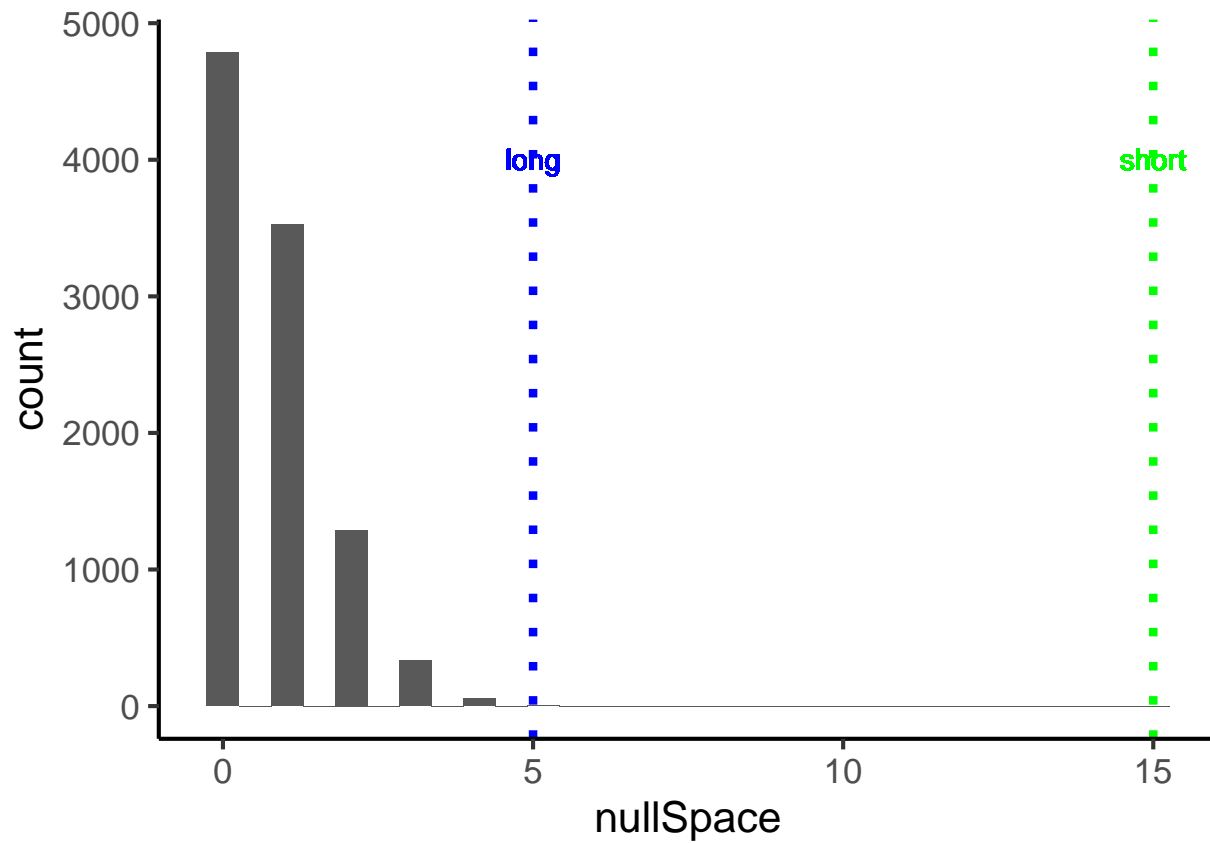
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



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

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

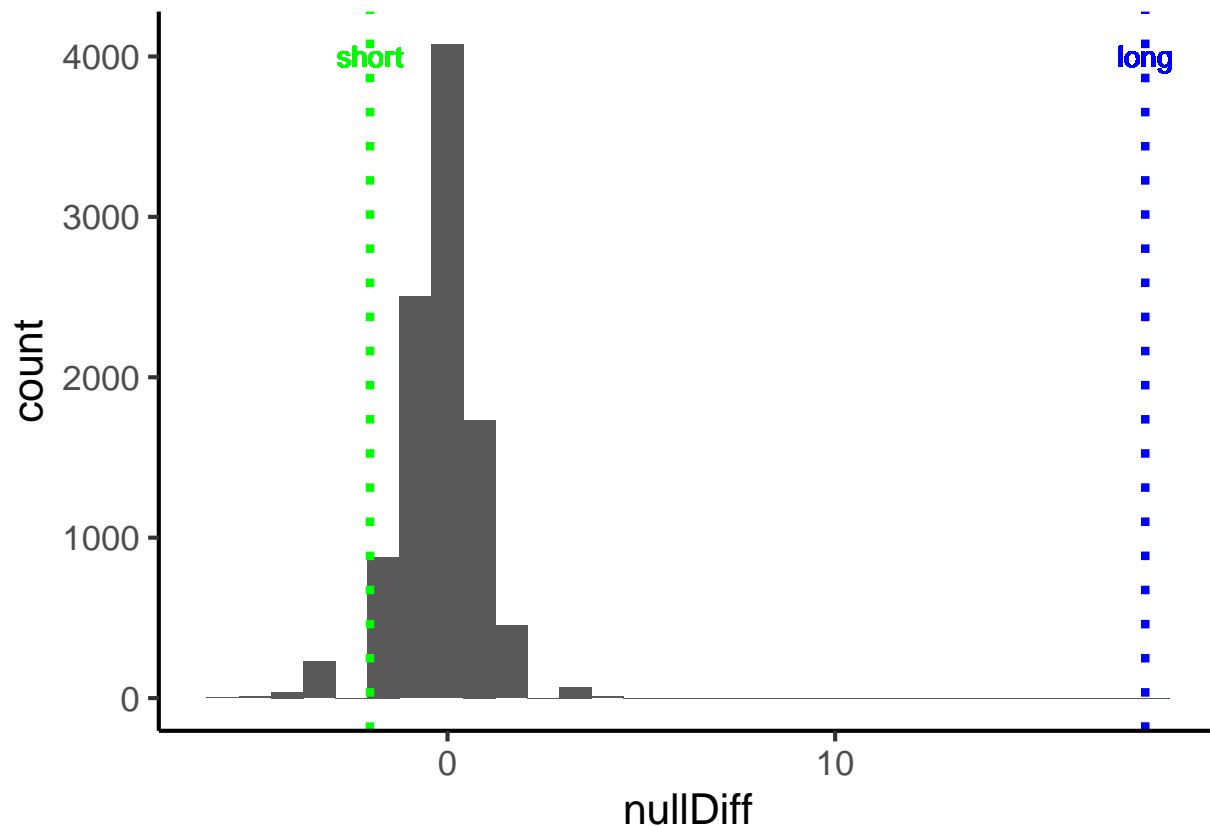
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



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

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

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



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 neighbors of long.

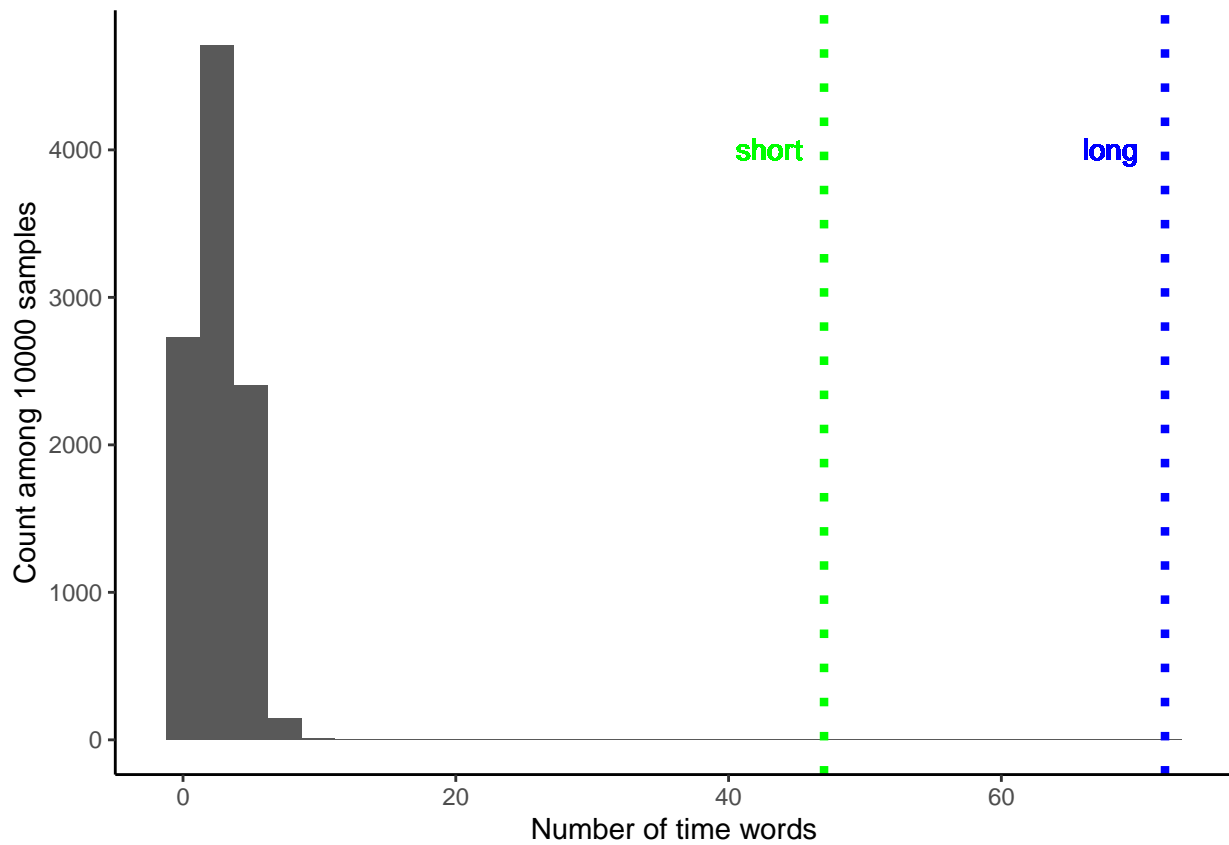
Below: Empirical values for the number of space words, the number of time words and the difference between the two within the first 500 nearest neighbors of short

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

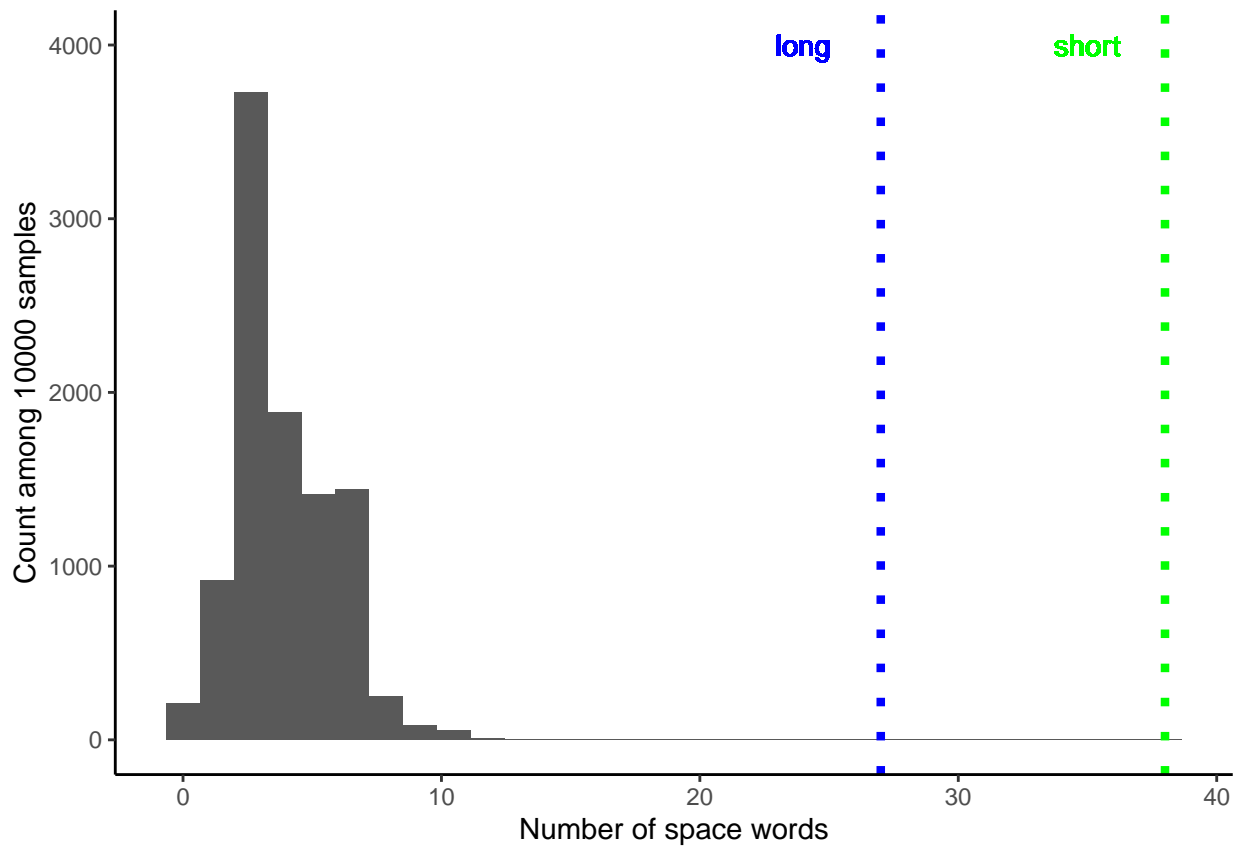
```
## Warning: Ignoring unknown parameters: text
```

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

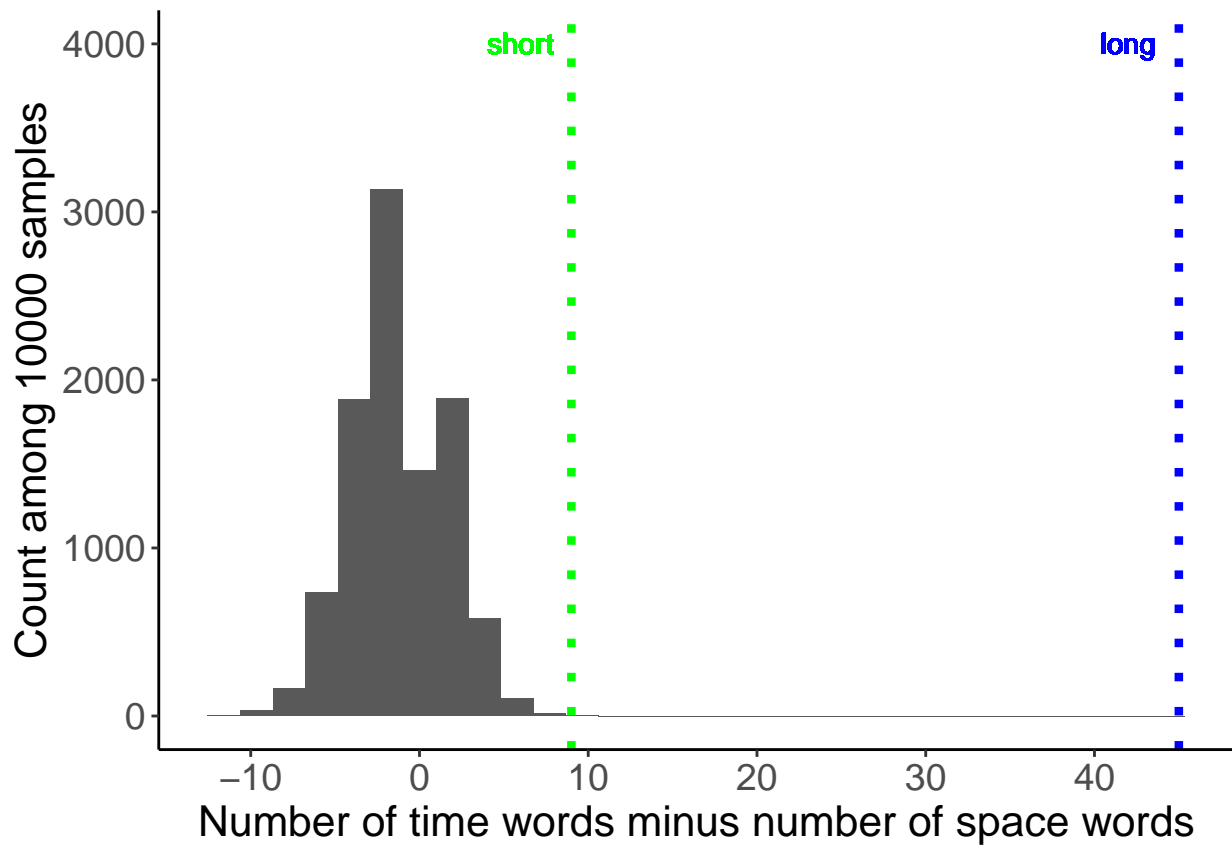
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Ignoring unknown parameters: text
## Warning: Ignoring unknown parameters: text
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
## Warning: Ignoring unknown parameters: text
## Warning: Ignoring unknown parameters: text
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Stats for long with samples N=500

For long

[1] 0

[1] 0

[1] 0

For short

[1] 1e-04

[1] 0

[1] 0