Systematicity over the course of early development: an analysis of phonological networks: Supplementary materials

Catherine E. Laing[1]

[1] University of York, York, UK

Author Note

All code and associated data for this manuscript can be found at https://github.com/cathelaing/NetworkGraphs. Data provided in the PhonBank corpus was collected through support from grant number NIH-NICHHD RO1-HD051698.

Correspondence concerning this article should be addressed to Catherine E. Laing, Department of Language and Linguistic Science, University of York, Heslington, YO10 5DD.. E-mail: catherine.laing@york.ac.uk

Systematicity over the course of early development: an analysis of phonological networks: Supplementary materials

S1: Testing model fit

Nested model comparisons were initially used to establish the best model fit for the data without the inclusion of Data type. Model comparisons testing Network size (*numNodes in the model*) are shown below. Including Network size alongside Age improved model fit for all models except the initial model testing mean path length.

Mean Path Length.

```
## Data: subset(SWD_red, data_type %in% c("actual", "WS_actual", "Erdos_Renyi"))
## Models:
## MPL.0: path_length ~ corpus + age + (1 + age | Speaker)
## MPL.1: path_length ~ corpus + age + numNodes + (1 + age | Speaker)
##       npar    AIC    BIC  logLik deviance Chisq Df Pr(>Chisq)
## MPL.0    7 1203.3 1231.8 -594.65   1189.3
## MPL.1    8 1205.6 1238.2 -594.81   1189.6     0  1          1
```

Clustering coefficient.

```
## Data: subset(SWD_red, data_type %in% c("actual", "WS_actual", "Erdos_Renyi"))
## Models:
## CC.0: clust_coef_avg ~ corpus + age + (1 + age | Speaker)
## CC.1: clust_coef_avg ~ corpus + age + numNodes + (1 + age | Speaker)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## CC.0    7 376.16 404.63 -181.08   362.16
## CC.1    8 369.98 402.53 -176.99   353.98 8.1732  1   0.004251 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Data type: Mean path length.

```
## Data: subset(SWD_red, data_type %in% c("actual", "target"))
## Models:
## MPL.DT.0: path_length ~ corpus + age + (1 + age | Speaker)
## MPL.DT.1: path_length ~ corpus + age + numNodes + (1 + age | Speaker)
##          npar     AIC     BIC logLik deviance  Chisq Df Pr(>Chisq)
## MPL.DT.0    7 -273.91 -248.15 143.96  -287.91
## MPL.DT.1    8 -278.53 -249.09 147.27  -294.53 6.6222  1    0.01007 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Data type: Clustering coefficient.

```
## Data: subset(SWD_red, data_type %in% c("actual", "target"))
## Models:
## CC.DT.0: clust_coef_avg ~ corpus + age + (1 + age | Speaker)
## CC.DT.1: clust_coef_avg ~ corpus + age + numNodes + (1 + age | Speaker)
##         npar     AIC     BIC logLik deviance  Chisq Df Pr(>Chisq)
## CC.DT.0    7 -632.59 -606.83  323.3  -646.59
## CC.DT.1    8 -664.80 -635.35  340.4  -680.80 34.205  1   4.96e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

S2: Running figures with Age instead of Network size

In the main paper, figures showing change in mean path length and clustering coefficient are visualised in relation to changing network size. Figures 1-4 plot the same data according to Age in months.

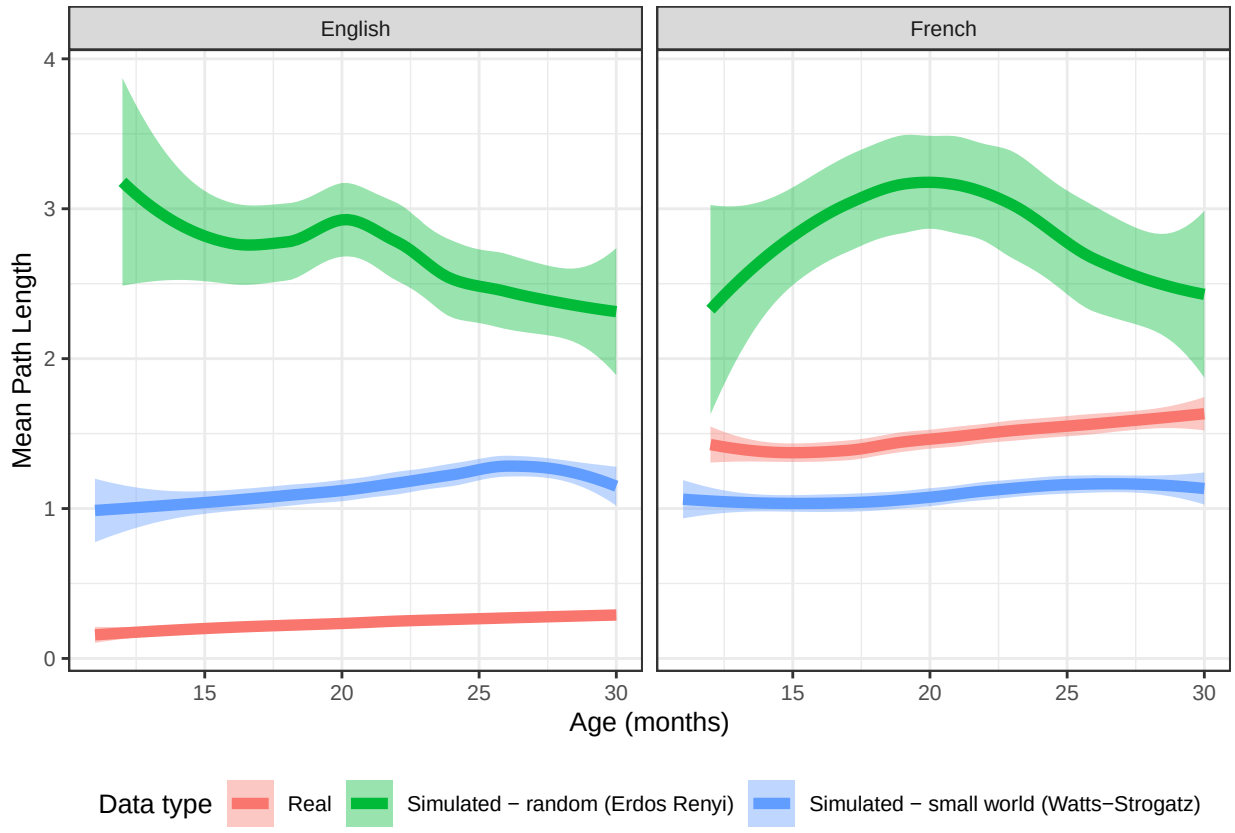

Figure 1. Change in mean path length over time, in Real data (Actual) compared with Simulated small-world (Watts-Strogatz) and random (Erdős–Rényi) networks. Coloured lines represent Data type; coloured bands represent 95% CIs. English and French data is plotted separately.

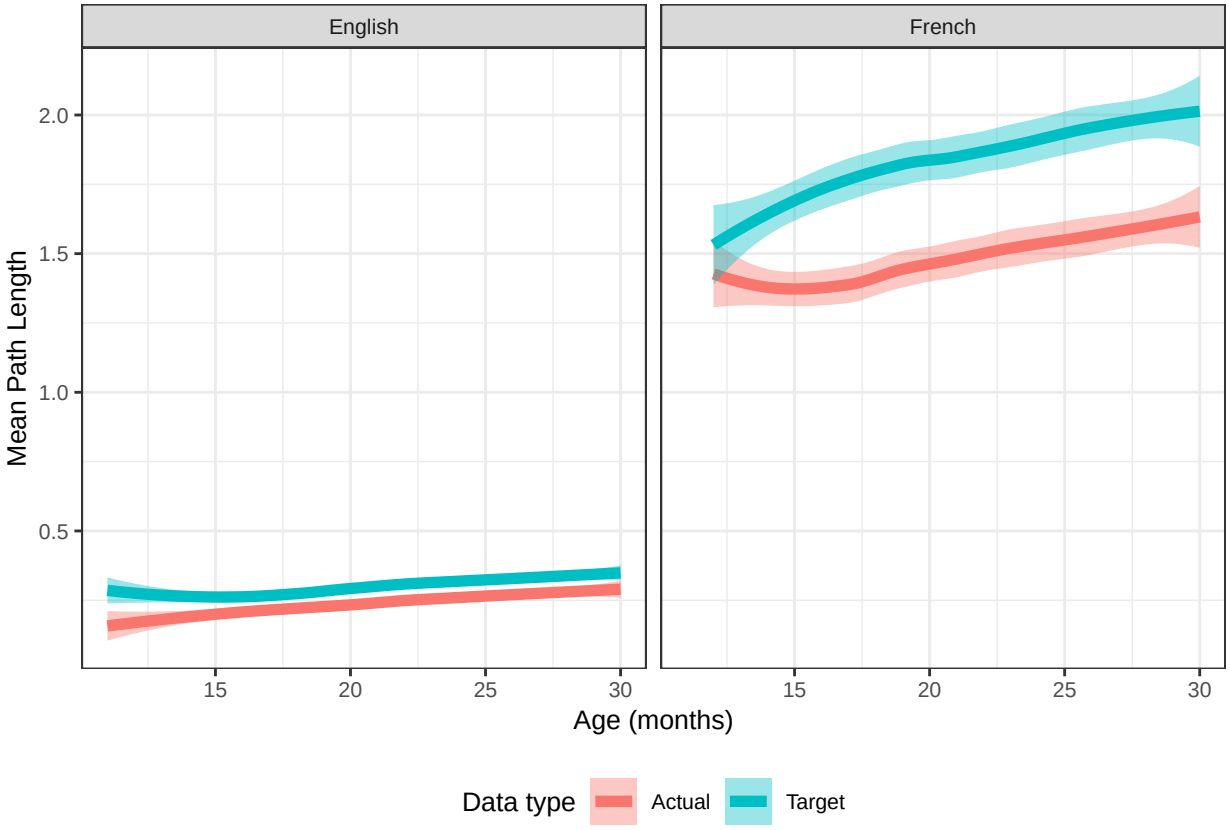

Figure 2. Change in mean path length over time, in Actual vs. Target data. Coloured lines represent Data type; coloured bands represent 95% CIs. English and French data is plotted separately.

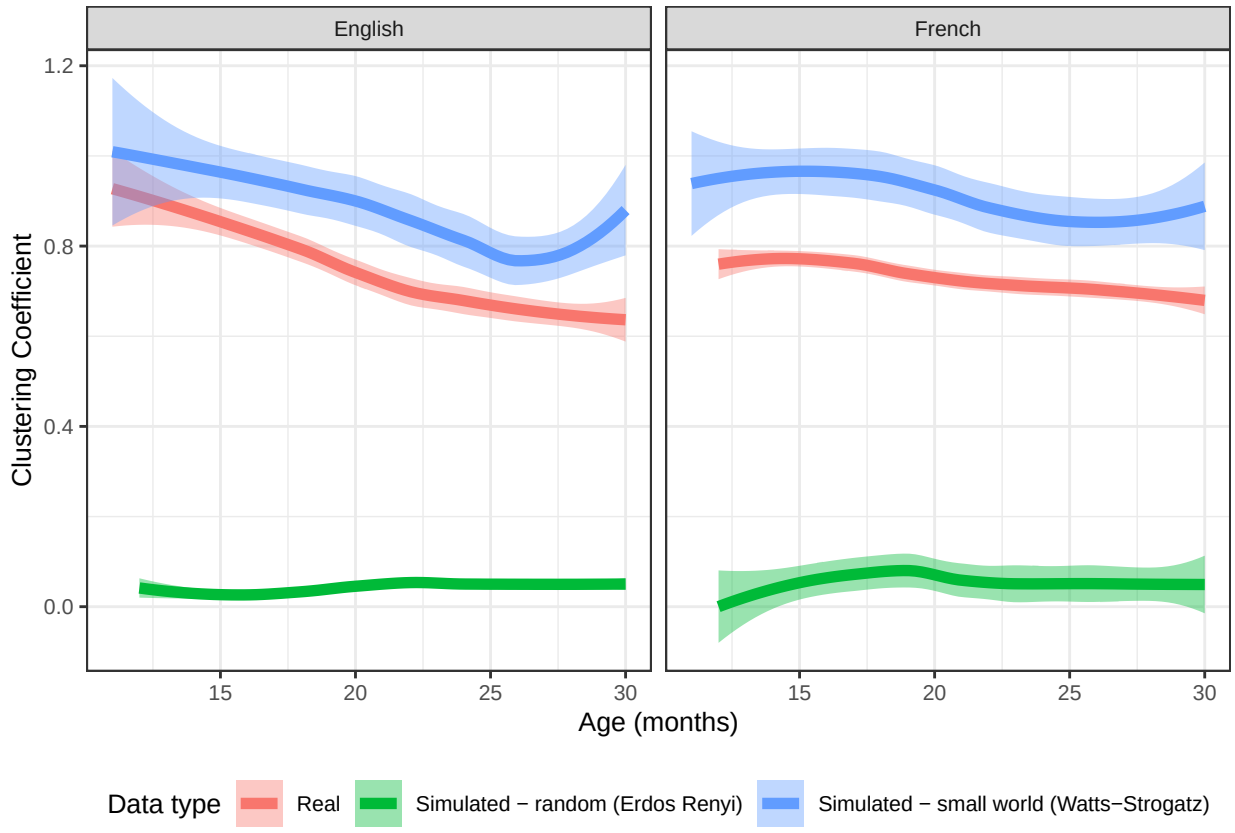

Figure 3. Change in mean clustering coefficient over time, in Real data (Actual) compared with Simulated small-world (Watts-Strogatz) and random (Erdős–Rényi) networks. Coloured lines represent Data type; coloured bands represent 95% CIs. English and French data is plotted separately.

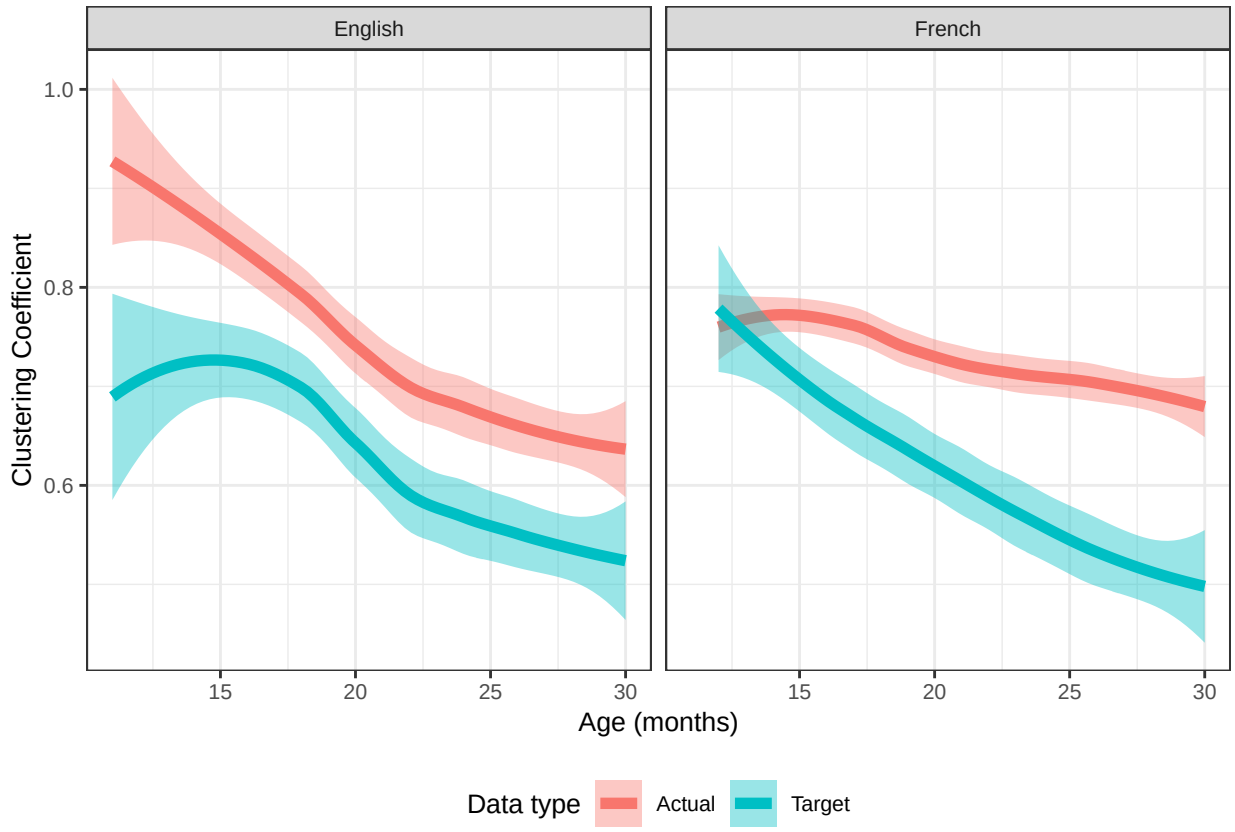

Figure 4. Change in mean clustering coefficient over time, in Actual vs. Target data. Coloured lines represent Data type; coloured bands represent 95% CIs. English and French data is plotted separately.