Github =

the markdown file is "Code for portfolio 2 part 1".

Describe the characteristics of the two groups of participants and whether the two groups are well matched.

The sample included 35 (excluding participant 66 because of missing data and only completing visit1) (28 male and 6 female) typical developing (TD) children with a mean age of 43±9 months (mean±sd) and 31 (26 male and 5 female) children suffering from Autistic syndrome Disorder (ASD) with a mean age of 30.6±7.23 months, a t-test showed that the two groups where not matched in age $t(325.71) = 14.41$, $p < 0.01$, (all t-tests reported are reported as ASD vs TD, furthermore we used welch t-test which does not have the assumption of normality). Both TD and children with ASD were mostly white. The TD scored a mean of 0.95±1.8 on the Autism Diagnostic Observation Schedul (ADOS) while the children with ASD scored a mean of 14.11±4.2, a t-test showed a clear difference between the two groups in ADOS score $t(216.81) = 39.99$, $p < 0.01$.
A verbal and non-verbal IQ test from the first visits showed that the TD had a verbal IQ of 20.14±5.1 and a non-verbal IQ of 25.92±3.39 while the children with ASD had a verbal IQ of 17.58±7.4 and a non-verbal IQ of 26.89±5.6. VerbalIQ significantly differed between the two groups $t(303,57)=-3.8$ $p < 0.01$, nonverbalIQ did significantly differ between groups $t(283.42)=1.98$ $p = 0.05$.
Socialization scores from the first visits showed that TD had a socialization score of 100.51±6.75 while the children with ASD had a socialization score of 77.20±9.63, a t-test revealed that socialization differed between the two groups $t(309,83)=-26.68$ $p<0.01$.
Lastly, we checked whether the two groups of children differed in their MLU, word tokens and word types in the first visit. Mean length of utterances for TD was 1.314±0.27 for ASD was 1.31±0.69, a t-test revealed that these values were not significantly different. The number of tokens used by the different groups did differ by a small margin $t(326.28)=-2.1617$ $p< 0.05$, but word type did not significantly differ between the groups.

From this we see that the children weren't matched for a lot of the variables, such as age, socialization score, verbalIQ scores, nonverbalIQ scores, but for the linguistic properties they seem to be quite well matched for instance in their MLU and word types, the significant difference in number of tokens could largely be due to the fact that we made a lot of tests without correcting for multiple comparison, if we did correct for that, using Bonferroni correcting our p-value threshold should have been 0.05*1/number of tests which in this case is 8, giving us a p-value threshold of 0.00625 if we want to keep our type 1 error rate at 5%. Doing this the verbal and non-verbal IQ-scores would also not be deemed significantly different. We therefore conclude that the children were well matched in the current study.

Hypothesis: The child's MLU changes: i) over time, ii) according to diagnosis

Let's start with a simple mixed effects linear model

Remember to plot the data first and then to run a statistical test.

See figure 0.

- Which variable(s) should be included as fixed factors?

Diagnosis for testing whether it effects children MLU changes. And visit to see if their MLU changes over time, one could include a lot of the other variables as fixed effects however due to our hypothesis we wouldn't include anything else.

- Which variable(s) should be included as random factors?

random slope for visit because some children will probably learn faster than others(we did not include this because the model failed to converge) and random intercepts for subject, because every child has a different starting point.

How would you evaluate whether the model is a good model?

First we evaluated the model by checking whether it met the assumptions of a linear mixed effect model.

No patten is to be observed of the residual-plot therefore linearity assumption is met.

The assumption of Homoskedasticity can also be checked from the residual-plot, it seems like there might be a bit of heteroskedasticity, but not too badly.

The output of the model shows the correlation between the two predictors to be 0.369, which is a small to medium correlation, no reason for concern for the assumption of multicollinearity.

normality of residuals assumption is also met by looking at the histogram or the qq-plot.

Then we estimated the R^2 marginal and conditional which were 0.349 and 0.77 respectively, showing that out fixed effects explains 35% of the variance and the whole model explains 77% of the variance. We think that, that is a decent good model.

Growth curve fitting:

- We then made 3 new models including quadratic, cubic and quartic aspects of visit, here we used the poly-function which takes care of visit being correlated with itself in the different polynomials, hereby not violated the assumption of multicollinearity.

- We then used the anova function to see which model was best, we assed this using the Bayesian information criterion(BIC) we used BIC because it has been shown to be more

conservative than AIC, and when doing these kinds of exploratory analyses we think that going with the more conservative is better.

The anova showed that a combination of linear and quadratic was the best model to the data which we then plotted in figure 1.

Let's check whether the model is doing an alright job at fitting the data. Plot the actual CHI_MLU data against the predictions of the model fitted(model)

- See figure 2. This plot shows that the model is pretty good at explaining the data, if all the points were on the straight line the model would be perfect at explaining the data (explaining 100% of the variance), however that doesn't tell us how good the model is at predicting new data.

Now it's time to report our results.

There was a significantly difference between ASD and TD in MLU B = 0.6472, se = 0.1636, t(58.8417) = 3.957, p<0.01

when the visit number increases linearly the MLU increases significantly aswell B=3.31, se = 0.6438 t(287.22) = 5.14, p < 0.01

when the visit number increases quadraticly the MLU decreases significantly B=-1.3, se = 0.64 t(287.24) = -2.1, p < 0.05

furthermore there was a interaction effect between diagnosis and visit (linearly) B=8.2 se = 0.89 t(287.27) = 9.2 p < 0.01

- A plain word description of the results

TD has longer length of utterances compared to ASD on average 0.6 longer utterances. The cildren's MLU increases as time passes linearly, however it seems to be the case that this linear increase decreases over time (it becomes less steep). However, these main effects should not be interpreted because of a significant interaction effect. The interaction effect shows that there is a difference in how much the MLU changes over time in the two groups, see figure 3, here its easy to see that the TD group's MLU increases as visits increase whereas the ASD group don't seem to increase their MLUs.

- A plot of your model's predictions (and some comments on whether the predictions are sensible)

a lot of the data falls outside the standard error which arguable makes it less sensible, see figure 4.

Hypothesis: Parental MLU changes: i) over time, ii) according to diagnosis

- We tested this hypothesis with a model including the mothers MLU which was predicted by Visit and Diagnosis. We saw that the mothers MLU changed over time and according to diagnosis, however diagnosis was barely significant, again if we corrected for multiple comparison using bonferroni's correction this would have been non-significant.

Parent MLU is affected by Diagnosis (ASD vs TD) B = 0.364, se = 0.15 t(147) = 2.44 $p < 0.05$ and visist B = 0.098, se = 0.20, t(289.6) = 5, $p < 0.01$. no interaction effect was found.

Your task now is to figure out how to best describe the children linguistic trajectory. The dataset contains a bunch of additional demographic, cognitive and clinical variables (e.g.verbal and non-verbal IQ). Try them out and identify the statistical models that best describes your data (that is, the children's MLU). Describe how you selected the best model and send the code to run the model to Victor and Byurakn.

we have not made models with all combinations, we made 2 overfitted models and compared them. The second model is better both in AIC and the explain variance for the fixed effects. This can be explained by the fact that the second model has more fixed effects, meaning that it explains the data better than the other models but doesnt nessarily mean it predicts new data better.

model 1 has an AIC value og 518 and a R^2 marignal 0.63

model 2 has an AIC value of 211 and a R^2 marignal 0.69