Supplementary Materials

Mixed effects models to investigate learning during test

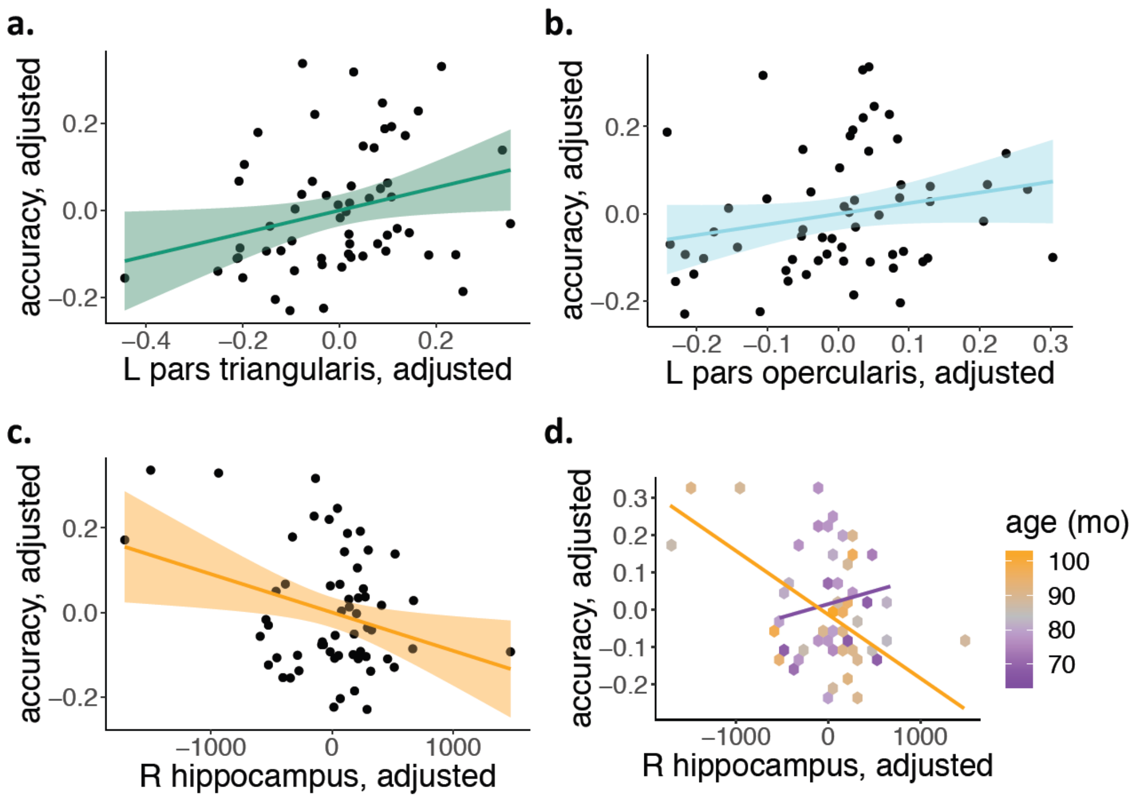
This model predicted accuracy and included word repetition (first, second or third occurrence, coded as 0, 1 and 2) as a fixed effect. Random intercepts and random slopes for the within-subject variable (word repetition) were included as fixed effects, grouped by subject:

glmer(accuracy ~ WordRep + (WordRep ||Subject), family=binomial)

We did not observe an effect of word repetition (β = 0.106, z = 1.07, p = .285).

Performance below 50% recoded as 50%

The performance of children who were below 50% were recoded as 50%, thus treating all learners at 50% or below as demonstrating the same amount of learning (no learning). In total 9 children were recoded from below 50% to 50%. As previously observed, thickness was positively associated with performance in the left pars triangualarus (β = .27, p = .0142) and marginally so in the left pars opercularis (β = .213, p = .112). Likewise, the right hippocampus was negatively associated with performance (β = -.322, p = .0213), and age interacted with hippocampal volume to predict performance (β = -4.62 p = .044) such that the association between hippocampal volume and task performance was stronger for older relative to younger children.

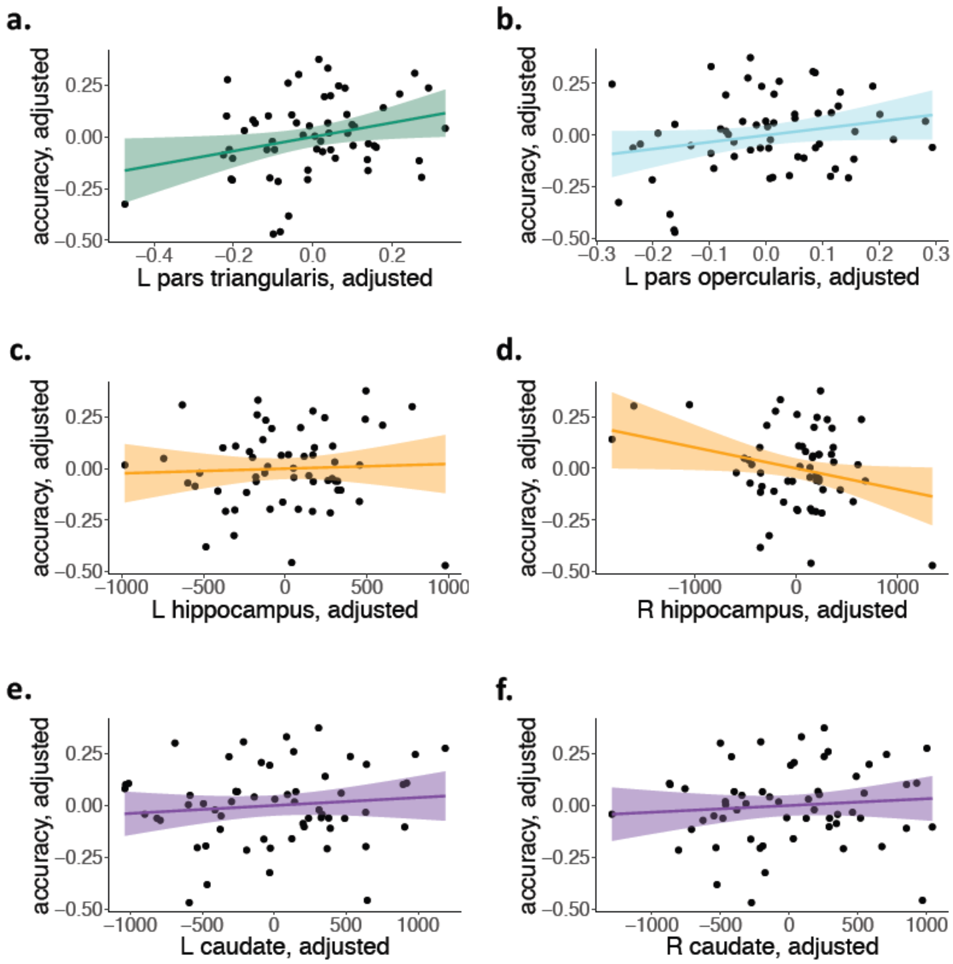


Supplementary Figure 1

Scatter plots depicting the relationship between adjusted performance (y-axis) in which below 50% was recoded as 50% and (a) adjusted left pars triangularis thickness (x-axis), (b) adjusted left pars opercularus thickness, and (c,d) right adjusted hippocampal volume. Solid lines depict the linear model and the opaque width of the line represents the 95% confidence interval. (d) Colored lines depict the regression between volume and performance separately for younger (purple) and older (orange) children as determined by a median split (at 83 months).

Structure Performance relationships not accounting for IQ

Thickness was positively associated with performance in the left pars triangualarus (β = .28, p = .01354) and marginally so in the left pars opercularis (β = .232, p = .0842). The right hippocampus was negatively associated with performance (β = -.285, p = .0481), but not the left hippocampus (β = .05, p = .743), nor the right or left caudate (right: β = .097, p = .493; left: β = .114, p = .415).



Supplementary Figure 2

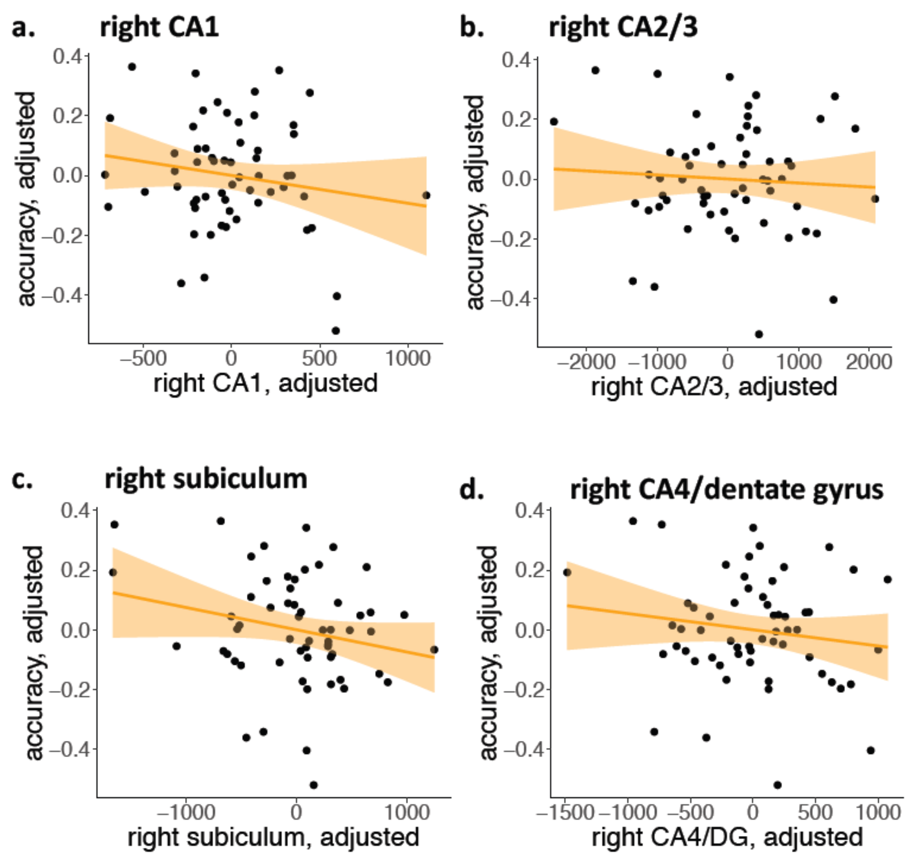
Scatter plots depicting the relationship between adjusted performance and adjusted (a) left pars triangularis thickness, (b) the left pars opercularus thickness, (c) left hippocampal volume, (d) right hippocampal volume, (e) left caudate volume, and (f) right caudate volume when models do not control for IQ. Solid lines depict the linear model and the opaque width of the line represents the 95% confidence interval.

Hippocampal subfield analyses

Method: Analyses were completed using FreeSurfer Version 5.3 (Fischl & Dale, 2000; Fischl et al., 2004) and the Segmentation of hippocampal subfields tools (https://surfer.nmr.mgh.harvard.edu/fswiki/HippocampalSubfields) (Iglesias et al., 2015).

Multiple linear regression analyses were performed to uncover the relationship between subfield volumes and measures and each child's mean performance on the statistical learning task. All models controlled for age, sex, IQ and total intracranial volume (ICV). Analyses were restricted to the right hemisphere and the following subfields of interest: CA1, CA2/3, subiculum, CA4/dentate gyrus.

<u>*Results*</u>*:* None of the investigated sub-regions were associated with performance on the statistical learning test: right CA1 (β = -.18, p =.215), right CA2/3 (β = -.07, p =.627), right subiculum (β = -.237, p =.0998), right CA4/dentate gyrus (β = -.155, p =.273), though note the trend in the right subiculum (p =.0998).



Supplementary Figure 3

Scatter plots depicting the relationship between adjusted performance and adjusted (a) right CA1 volume, (b) rightCA2/3 volume, (c) right subiculum volume, and (d) right CA4/DG volume. Solid lines depict the linear model and the opaque width of the line represents the 95% confidence interval.

Control analysis: thickness in primary visual cortex

To investigate whether the relationship between thickness and performance on the statistical learning testis specific to our predicted cortical regions in the inferior frontal cortex, we ran an analysis with a control region that we believe should *not* be involved in auditory statistical learning in any direct way: primary visual cortex (the cuneus), also in the left hemisphere. As expected, there was no relationship between the thickness of this region and performance on the statistical learning measure (β = -.001, p = .945). As in the analyses reported in the main text, this model controlled for age, sex, IQ and total intracranial volume (ICV).

References

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A, 97*(20), 11050-11055.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cereb Cortex, 14*(1), 11-22.

Iglesias, J. E., Augustinack, J. C., Nguyen, K., Player, C. M., Player, A., Wright, M., et al. (2015). A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *Neuroimage, 115*, 117-137.