

The relationship between brain structure and function during novel grammar learning across development

Merel E.E. Koning^{1,†}, Nina K. Wyman^{1,†}, Willeke M. Menks^{1,3}, Clara Ekerdt¹, Guillén Fernández¹, Evan Kidd^{3,4,5}, Kristin Lemhöfer¹, James M. McQueen^{1,3}, Gabriele Janzen^{1,2,*}

¹Donders Institute for Brain, Cognition and Behaviour, Radboud University and Radboud University Medical Center, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

²Behavioural Science Institute, Radboud University, Thomas van Aquinostraat 4, 6525 GD Nijmegen, The Netherlands

³Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

⁴ARC Centre of Excellence for the Dynamics of Language, Australian National University, Building 9, H.C. Coombs Building The, ACT 2601, Canberra, Australia

⁵School of Literature, Languages, and Linguistics, Australian National University, 110 Ellery Crescent Acton, ACT 2601, Canberra, Australia

*Corresponding author: Donders Institute for Brain, Cognition and Behaviour, Radboud University and Radboud University Medical Center, Thomas van Aquinostraat 4, 6525 GD Nijmegen, The Netherlands. Email: gabriele.janzen@ru.nl

†Merel E.E. Koning and Nina K. Wyman contributed equally to this work.

In this study, we explored the relationship between developmental differences in gray matter structure and grammar learning ability in 159 Dutch-speaking individuals (8 to 25 yr). The data were collected as part of a recent large-scale functional MRI study (Menks WM, Ekerdt C, Lemhöfer K, Kidd E, Fernández G, McQueen JM, Janzen G. Developmental changes in brain activation during novel grammar learning in 8–25-year-olds. *Dev Cogn Neurosci*. 2024;66:101347. <https://doi.org/10.1016/j.dcn.2024.101347>) in which participants implicitly learned Icelandic morphosyntactic rules and performed a grammaticality judgment task in the scanner. Behaviorally, Menks et al. (2024) showed that grammaticality judgment task performance increased steadily from 8 to 15.4 yr, after which age had no further effect. We show in the current study that this age-related grammaticality judgment task performance was negatively related to cortical gray matter volume and cortical thickness in many clusters throughout the brain. Hippocampal volume was positively related to age-related grammaticality judgment task performance and L2 (English) vocabulary knowledge. Furthermore, we found that grammaticality judgment task performance, L2 grammar proficiency, and L2 vocabulary knowledge were positively related to gray matter maturation within parietal regions, overlapping with the functional MRI clusters that were reported previously in Menks et al. (2024) and which showed increased brain activation in relation to grammar learning. We propose that this overlap in functional and structural results indicates that brain maturation in parietal regions plays an important role in second language learning.

Key words: brain maturation; cortical thickness; grammar learning; gray matter volume; second language acquisition.

Introduction

The age at which people start to learn a second language (L2) is an important factor for L2 learning success; the path toward proficiency is typically easier and more successful when people start learning a second language early in childhood, with the ability to master some components of language like phonology and grammar declining with age (e.g. Johnson and Newport 1989; Flege 1995; Hartshorne et al. 2018). Although the exact timing of this decline in language learning ability remains uncertain, it seems to occur sometime between childhood and adolescence. For grammar learning—the focus of this paper—a decline in learning ability has been suggested to occur sometime between middle childhood and late adolescence (Johnson and Newport 1989; Hartshorne et al. 2018). In order to uncover the neural mechanisms potentially responsible for this developmental shift, our group recently reported on a large-scale functional magnetic resonance imaging (fMRI) study that investigated brain activation during novel grammar learning in 165 participants aged between 8 and 25 yr (Menks et al. 2024). Following up on those fMRI results and using the structural MRI data from the same developmental sample, the aim of the current study was to examine how differences in gray matter structure across development relate to novel

grammar learning and whether such developmental differences can be related to the fMRI results.

In the study of Menks et al. (2024), participants, who were unfamiliar with the Icelandic language, learned Icelandic morphosyntactic rules over the course of a week and performed a grammaticality judgment task (GJT) in the MRI scanner in order to allow investigation of developmental differences in brain activation linked to novel grammar learning. Menks et al. (2024) found age effects in performance on the GJT at the behavioral and neurological level. Behaviorally, GJT performance was positively related to age until 15.4 yr of age, after which age had no effect on grammar learning performance anymore. This means, paradoxically, that although in general children ultimately obtain a higher proficiency when learning a new language, adults outperform children in the initial stages of language learning, a finding which is in line with the second language acquisition literature (e.g. Asher and Price 1967; Snow and Hoefnagel-Hohle 1978; Krashen et al. 1979) and findings from more recent studies (Tatsuno and Sakai 2005; Ferman and Karni 2010; McNealy et al. 2011; Hashizume et al. 2014; Takashima et al. 2019). Neurally, Menks et al. (2024) found that these age-related differences in grammar learning were related to activation levels in

working-memory and grammar-related brain areas, including the bilateral inferior frontal gyrus, the fusiform gyrus, the right inferior temporal gyrus, the right superior parietal gyrus, and the bilateral inferior parietal gyri. Although adults and children activated similar brain regions, activation level was age dependent, with children showing less activation than older participants. This suggests that children and adults recruit the same networks during novel grammar learning, but they do not recruit them to the same degree. Thus, the fMRI results showed that there is a developmental change in the degree, but not the location of brain activation in the initial stages of grammar learning.

Structural brain development

Besides changes in brain activation, the brain also undergoes structural changes during development. Early studies found that the development of gray matter volume and cortical thickness follows an inverted U-shaped trajectory; after an initial increase, gray matter volume and cortical thickness peak in late childhood or early adolescence, after which they decrease again (Giedd et al. 1999; Gogtay et al. 2004; Shaw et al. 2008; Raznahan et al. 2011). This decrease is thought to be associated with synaptic pruning and increased myelination resulting in increased efficiency (Sowell et al. 2004a). More recent studies suggest that gray matter decrease starts much earlier than originally assumed, and observed a decline already at 8 yr of age or even earlier (Sowell et al. 2004b; Tamnes et al. 2010; Muftuler et al. 2011; Koolschijn and Crone 2013; Wierenga et al. 2014; Fjell et al. 2015; Mills et al. 2016; Vidal-Pineiro et al. 2020). Bethlehem et al. (2022) recently aggregated 123,984 scans from >100 studies and found gray matter volume peaking at 5.9 yr followed by a near-linear decrease. Cortical thickness was found to be peaking already as early as 1.7 yr of age. Although the exact timing of the onset of the decline remains uncertain, it is likely that findings from recent studies that observed a decrease already in early and middle childhood, which continues in late childhood, adolescence, and beyond, is more accurate, considering methodological advances and improved data quality (Norbom et al. 2021).

Gray matter decline is seen throughout the cortex across development, but regional differences in the rate and timing of this general pattern have been observed (Giedd et al. 1999; Sowell et al. 2001; Gogtay et al. 2004; Bethlehem et al. 2022). Overall, studies suggest that gray matter development follows a posterior-to-anterior or sensory-to-association cortices gradient of maturation, which means that primary sensory areas mature earlier than the higher order association cortices (Gogtay et al. 2004; Shaw et al. 2008; Walhovd et al. 2017; Norbom et al. 2021; Bethlehem et al. 2022). Furthermore, variable trajectories have been observed for dorsal frontal and parietal regions (Sowell et al. 1999a; Sowell et al. 1999b; Sowell et al. 2001; Sowell et al. 2003); in parietal cortices, gray matter seems to decline most rapidly between childhood and adolescence (Sowell et al. 1999b) and in frontal lobes between adolescence and young adulthood (Sowell et al. 1999b). This delayed maturation of frontal structures parallels the continued development of executive functions between adolescence and adulthood supported by these structures (Fuster 2002).

Although maturational trajectories are generally associated with gray matter reduction, cortical thickening has also been observed across development, for example, in relation to skill learning or training. In studies of children's first language development, sentence comprehension (Fengler et al. 2015), receptive and expressive language ability (Deniz Can et al. 2013), phonological processing (Lu et al. 2007), and vocabulary knowledge (Richardson et al. 2010) have been associated with gray matter

increases in frontal and temporal language regions and the right hippocampus. In adults, L2 learning has also been related to gray matter increases (Wong et al. 2008; Mårtensson et al. 2012; Hosoda et al. 2013). Mårtensson et al. (2012) observed increases in right hippocampal volume and in cortical thickness in the left middle frontal gyrus (MFG), the left inferior frontal gyrus (IFG), and the left superior temporal gyrus (STG) after 3 mo of intense foreign language training. Furthermore, proficiency obtained in the foreign language was related to an increase in right hippocampal volume and left STG cortical thickness. In children aged 5 to 8, Finn et al. (2019) investigated statistical learning in an artificial language. They found that thickness of the left IFG and volume of the right hippocampus predicted statistical learning ability. Apart from gray matter increases, decreases as a function of language experience have also been observed. Kepinska et al. (2023a) examined in adult participants the impact of multilingualism on the morphology of the auditory cortex and found that language experience was negatively related to cortical thickness of the second transverse temporal gyrus. Taken together, these studies show that language learning can be related to changes in gray matter structure in both children and adults. However, it has yet to be examined how novel grammar learning is related to changes in gray matter structure across development.

Structure and function overlap

Changes in brain structure and function co-occur during development and investigating how both structural and functional maturation relate to novel grammar learning might give us more insight into the neural correlates involved. Many studies have combined different neuroimaging techniques in order to investigate the complex relationship between brain function and structure in relation to behavioral measures such as memory, attention, executive control, arithmetic, and cognitive decline (e.g. Bangen et al. 2012; Chavan et al. 2015; Gur et al. 2021; Hu et al. 2014; Lerman-Sinkoff et al. 2017; Persson et al. 2006; Rivera et al. 2005; Salami et al. 2014; Squeglia et al. 2013; Sun et al. 2016; Tavakol et al. 2021; for an overview see Litwińczuk et al. 2023). The development of language skills has also been related to overlapping structural and functional maturation. Lu et al. (2009) identified brain regions where functional activation during orthographic processing was related to cortical thickness. They found that a "mature" pattern of activation (i.e. activation intensity correlated with skill level improvement) was associated with thinner cortex in the right frontoparietal region in children from 6 to 15 yr. Nuñez et al. (2011) observed that thickness in inferior frontal regions involved in syntax processing was negatively related to brain activation in children aged 7 to 15. Richardson et al. (2010) found that both gray matter density and brain activation during sentence processing correlated with vocabulary knowledge in left posterior temporal regions throughout the lifespan. These studies suggest that there is a relationship between gray matter structure and brain function, but no study has examined this in relation to novel grammar learning.

The present study

The aim of the current study was to investigate how developmental differences in gray matter structure relate to differences in brain activity during novel grammar learning in a large developmental sample with participants aged between 8 and 25 yr. To investigate this, we first examined which brain regions showed a correlation between age-related differences in novel grammar learning performance and gray matter structure (i.e.

volume and thickness). We expected to find that most brain regions would follow the general developmental trend of gray matter decrease. Therefore, since novel grammar learning performance is partly predicted by age, we expected to find negative correlations between gray matter structure and novel grammar learning performance. Next, we compared these structural results with the fMRI clusters reported in Menks et al. (2024), in order to find regions where both structural and functional maturation are related to novel grammar learning performance. Specifically, we expected to observe an overlap between regions in which we found a negative correlation between gray matter structure and age-related novel grammar learning performance and regions that showed age-related functional activation during the GJT, such as the bilateral inferior frontal gyri, the fusiform gyrus, the right inferior temporal gyrus, the right superior parietal gyrus, and the bilateral inferior parietal gyri, as reported in Menks et al. (2024). Finally, we performed additional analyses in order to further examine whether the overlapping regions are specifically related to novel grammar learning or whether they are also related to other language learning measures or more general supportive or memory functions. Therefore, we additionally investigated the relationship between gray matter structure and L1 and L2 grammar proficiency, L2 vocabulary knowledge, and memory abilities. These results were then again compared with the fMRI clusters.

Materials and methods

Participants

One hundred and ninety-five participants were recruited for the study. Adults were recruited using posters, social media, and the research participation system of Radboud University. Children were recruited through posters at schools, libraries, and community events. Participants were raised monolingually, with Dutch as their first language (L1). All were right handed, had normal or corrected-to-normal vision, and no hearing impairment or reading disorder. Other reasons for exclusion were a bilingual upbringing; a history of neurological, psychiatric, or cognitive disorders; or incompatibility with MRI. Thirty-six participants were excluded due to scanner malfunction ($n=5$), an incomplete dataset ($n=7$), reading difficulties ($n=2$), a late autism diagnosis ($n=1$), an incidental finding ($n=1$), artifacts in the scan ($n=1$), poor fMRI data quality (i.e. either poor fMRI scan quality or poor GJT performance) ($n=11$), and artifacts related to excessive head motion in the anatomical scans ($n=8$). This resulted in an overall inclusion of 159 participants (aged 8.26 to 25.97 yr, $M=18.10$, $SD=4.64$, 113 females, see Fig. S1 for the participant distribution). An a priori power analysis was conducted using G*Power (version 3.1.9.7; Faul et al. 2009) to determine the minimum sample size required to test the study hypothesis. Results indicated that the required sample size with 80% power for detecting a medium effect, at a significance criterion of $\alpha=0.05$, was $n=114$ for a linear multiple regression analysis with three predictors. Our obtained sample size is larger and therefore adequate to test the study hypothesis. Participants completed a language background questionnaire about additional languages they learned aside from Dutch. All participants were learning/spoke English as a second language and most participants reported to have learned other languages in addition to English, among which the most commonly named were German, French, and Spanish. However, none of the participants reported to have prior knowledge of Icelandic—the target language in the novel grammar learning task—before the experiment. Written informed consent was obtained from all participants and their parents (if participants were younger than

18) before participation. The study was approved by the ethical committee CMO region Arnhem–Nijmegen (2018-4561; 2014-288).

Procedure

The data of this study were acquired as part of a larger project investigating the neural correlates of age-related and individual differences in grammar learning ability. A detailed description of the study protocol can be found in Menks et al. (2022). In the project, participants took part in an Icelandic grammar training in which they learned a subset of Icelandic morphosyntactic rules. Participants completed three testing sessions. The first session was a behavioral session in which participants completed an extensive behavioral test battery, which included behavioral measures of language proficiency in the L1 (Dutch) and L2 (English), and memory abilities. The second and third session were MRI sessions in which participants did the Icelandic grammar training and a GJT in the scanner. Between the MRI sessions, participants continued with the Icelandic grammar training at home for 5 d.

Icelandic grammar training. Before starting the Icelandic grammar training, participants performed a word familiarization task (Fig. 1A). In this task, they learned 20 Icelandic nouns that were later used in the grammar training. The Icelandic nouns were all Dutch cognates (e.g. the Icelandic word *api* resembles the Dutch word *aap*, both meaning “monkey”). This ensured that the Icelandic nouns were easy to learn and that later GJT performance would reflect grammatical learning independent of lexical learning. The nouns were learned through a repeating word–picture matching task until all items were answered correctly. This word familiarization task was done twice on separate days before the start of the first MRI session. In the first MRI session, participants did the Icelandic grammar training in which they implicitly learned Icelandic morphosyntactic rules through an image–sentence matching game (Fig. 1B). The rules that were learned concerned inflection of word phrases based on gender (masculine/feminine), number (singular/plural), and case (nominative/accusative). Feedback was given after every trial, irrespective of whether the given answer was correct or incorrect. However, no explicit instructions regarding the grammar rules were given during the task. The task was adaptive and participants could end up in one of three levels of difficulty based on their performance. These three different levels were tested with several pilot studies, which ensured that the training was not too difficult for children or too easy for adults, thus avoiding ceiling and floor effects (Fig. S1 for the participant distribution per level and age bin). Between the two MRI sessions, participants completed an at-home Icelandic grammar training for 5 d. Each home training day was similar to the grammar training they did during the first MRI session, except that participants remained at the same level they achieved during the training in the first MRI session (for more details, see Menks et al. 2022).

fMRI task. In both MRI sessions, participants performed a GJT in the MRI scanner at the same level that was reached during the training. The task consisted of a baseline (25%) and a grammar condition (75%) with 192 trials. During the baseline, two real or scrambled Icelandic words (i.e. without sentence context) were shown, and participants had to indicate whether they had learned the word or not. The grammar condition consisted of grammatical (66%) or ungrammatical (33%) sentences. Participants judged whether the displayed sentences were correct or incorrect according to the implicitly learned Icelandic grammar rules. During the first MRI session, the GJT trials consisted of a white fixation cross, followed by an Icelandic sentence in black letters on a light-gray screen for 3,500 ms and ended with another fixation cross for

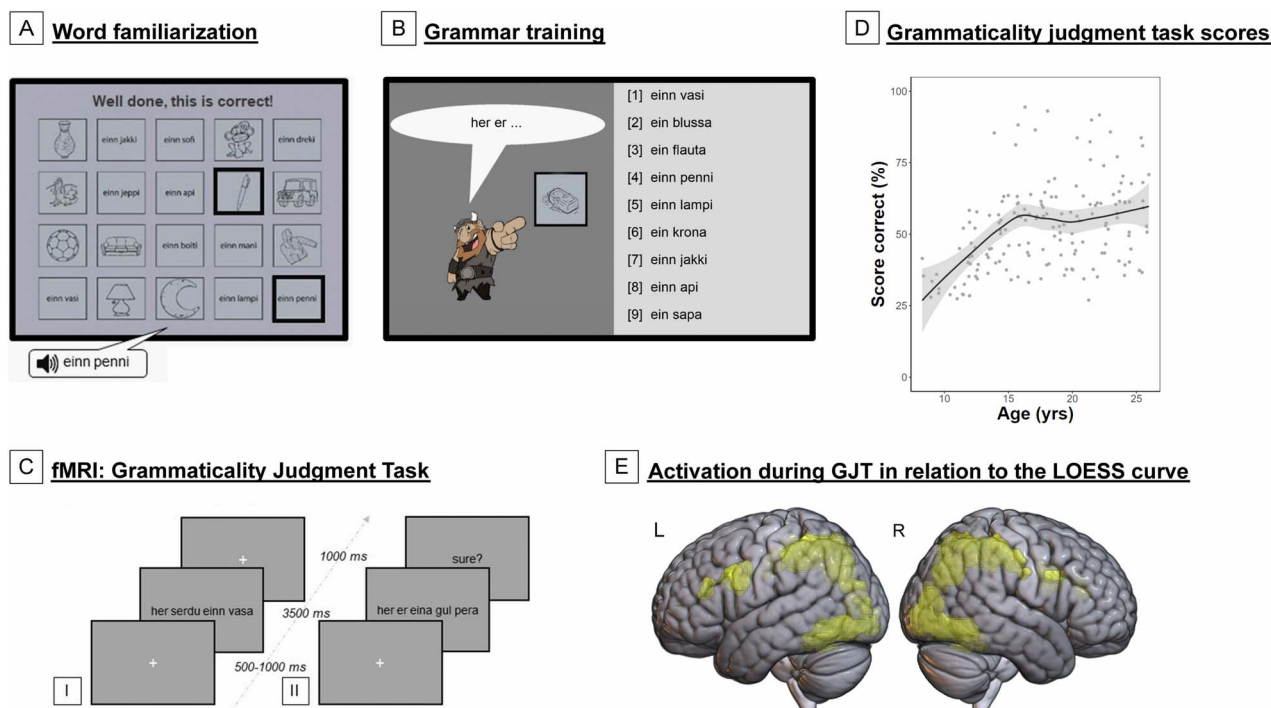


Fig. 1. Overview of the procedure and main results reported in Menks et al. (2024; adapted version of Figs. 1, 2, and 3). (A) The word familiarization task. (B) The Icelandic grammar training. (C) The fMRI GJT task before (I) and after (II) the 5-d home training. (D) Mean GJT scores (average of both MRI sessions) are plotted against age and overlaid with the LOESS curve (black). The shaded areas indicate the confidence interval (95%) of each estimated curve. (E) Whole-brain activation (cluster level, $P < 0.05$ FWE-corrected) during the GJT in relation to the LOESS fitted curve (i.e. the estimated combined effect of age and grammar learning performance).

1,000 ms. In the second MRI session, the GJT trials were identical to the first session, except that at the end of each trial the participant indicated how certain they were of their response, replacing the last 1000-ms fixation cross of the trial (Fig. 1C).

Behavioral measures

Language proficiency measures

L1 and L2 language proficiency was measured by three standardized tests, which included the recently developed Syntest (Janssen et al. unpublished) and the Test for Reception of Grammar (second edition, TROG-2) (Bishop 2009) as measures for L1 and L2 grammatical ability, respectively. The Peabody Picture Vocabulary Test fourth edition (PPVT-4) (Dunn and Dunn 2007) was used as a measure for L2 vocabulary knowledge. The raw scores for each test consisted of the total number of correct answers. From the raw score, a percentage was calculated by dividing the raw score by the maximum number of trials of the test. The raw scores (expressed as a percentage) were then used for further statistical analyses.

Memory abilities

Phonological short-term memory was measured using the nonword repetition task adapted from de Vos et al. (2019). Verbal working memory was measured using the digit span test, which included the forward, backward, and sorting task from the Wechsler Scale of Intelligence for Adults (WAIS; Wechsler 2008). Adults performed the WAIS-IV_NL version and participants under the age of 17 performed the Wechsler Scale of Intelligence for Children (Wechsler 2018). Declarative memory was measured using the word pairs test from the Dutch version of the Wechsler Memory Scale-IV (Bouman et al. 2012). For all memory ability tests, a total raw score was calculated. The total score of both versions

of the digit span consisted of the sum of correct answers on the forward, backward, and sorted digit span. These raw scores were then transformed into standardized scores. The total score of the nonword repetition task consisted of the total number of correctly repeated nonword sequences. The total score of the word pairs test consisted of the sum of the reproduction and recognition sections of the test. The standardized scores on the digit span and the total raw scores of the nonword repetition task and word pairs test were used for further statistical analyses.

Grammar learning performance

Performance on the GJT was analyzed previously (Menks et al. 2024). Using RStudio (R-4.1.3), a mean GJT score from the first and second MRI session was calculated for each participant and used as a measurement of overall grammar learning success. To examine age-related grammar learning performance, a LOESS—a local polynomial regression-smoothing procedure—curve-fitting function was applied to assess visually the model-free relationship between age and grammar learning performance (degree = 2, span = 0.75, method = “loess”, parametric = FALSE, drop.square = FALSE, normalize = TRUE) (Fig. 1D). This makes it possible to identify data patterns that may be overlooked with curve-fitting procedures that assume an a priori shape (Cleveland and Devlin 1988). The predicted values from the LOESS curve (i.e. the predicted grammar performance based on age) were used previously by Menks et al. (2024) as a measure of age-related grammar learning performance. We used this same measure in our statistical analyses.

Structural MRI data acquisition

Data were acquired on a 3-T Siemens SKYRA MRI scanner using a 32-channel head coil at the Donders Institute for Brain, Cognition and Behaviour in Nijmegen. Two whole-brain T1-weighted

(T1w) images were collected during two MRI sessions. One structural T1w image was acquired using a magnetization prepared rapid gradient echo (MPRAGE) sequence. The parameters for this MPRAGE sequence were as follows: repetition time (TR) = 2,000 ms; echo time (TE) = 2.01 ms; matrix size = 256 × 256; field-of-view (FOV) = 256 mm; flip angle = 8°; voxel size = 1 mm; slice thickness = 1 mm, 192 sagittal slices covering the entire brain. Parallel imaging (iPAT = 2) was used to accelerate the acquisition, resulting in an acquisition time of 4 min and 40 s. The second type of whole-brain T1w image acquired in this study was the MP2RAGE (Marques et al. 2010). The MP2RAGE sequence parameters were as follows: TR/TI1/TI2 = 5,000/700/2,500 ms; matrix size = 256 × 216; FOV = 256 mm; flip angle1 = 6°; flip angle2 = 5°; voxel size = 1 mm; slice thickness = 1 mm, 224 sagittal slices covering the entire brain. Parallel imaging (iPAT = 4.6) was used to accelerate the acquisition, resulting in an acquisition time of 4 min. Afterwards, a corresponding low-resolution B1 fieldmap was collected to correct for small inhomogeneities (Marques and Gruetter 2013). A total of 42 sagittal slices were acquired, covering the entire brain in an acquisition time of 20 s. The B1 fieldmap sequence parameters were as follows: TR = 1,000 ms; TE = 2.07 ms; matrix size = 256 × 216; FOV = 256 mm; voxel size = 1 mm; slice thickness = 1 mm. In order to minimize head movements, a small pillow was placed in the head coil to stabilize the participant's head and make them comfortable. Furthermore, a small skin-friendly piece of tape was placed across the forehead of each participant and the head coil base to provide them with head motion feedback, which is a child-friendly way to prevent head movement. Participants watched a short movie during the scans and were instructed to relax and lie as still as possible. The MP2RAGE was acquired during the first MRI session and the MPRAGE was acquired during the second MRI session. The aim was to collect the two types of T1w images for each participant. However, for some participants (n = 22) two MPRAGES were collected and for one participant only one T1w image was collected (MPRAGE).

Structural MRI preprocessing

If available, two T1w images were combined using fMRIPrep (v22.0.1). Voxel-based morphometry was then performed on (combined) T1w images using Cat12 (version 8.2, <http://dbm.neuro.uni-jena.de/cat/>). The images were normalized to a template space using Geodesic Shooting registrations (Ashburner and Friston 2011) and segmented into gray matter, white matter (WM), and cerebrospinal fluid. The segmentation outputs were manually assessed for overt segmentation failure and sample homogeneity. For each participant, the total intracranial volume (TIV) was estimated and the images were smoothed with a 6-mm full width at half maximum Gaussian kernel. Gray matter volume and cortical thickness output files were used for further analyses.

Statistical analyses

In the first main analysis, the relationship between gray matter structure and age-related novel grammar learning performance was examined by performing a second-level whole-brain regression analysis with TIV as a nuisance covariate. Values extracted from the LOESS curve, representing the predicted grammar performance based on age, were used as a predictor of gray matter volume and cortical thickness in these analyses. All results were thresholded using a cluster level correction of $P = 0.001$ and a family-wise error (FWE) correction of $P < 0.05$. The second main analysis explored how the structural results (i.e. the regions where we found a relationship between gray matter structure and age-related GJT performance) relate to the fMRI clusters reported

previously (Menks et al. 2024). To this end, a mask was created of the fMRI clusters and overlaid with our structural results. Overlapping clusters were then thresholded at a voxel level of $P = 0.001$ and FWE-corrected at $P < 0.05$ to determine the clusters with significant overlap.

Besides these main analyses, additional analyses were executed to investigate further relationships with gray matter structure. Firstly, to explore the relation between age and gray matter development in our sample, a whole-brain regression analysis was performed with age as a predictor of gray matter volume and cortical thickness. Secondly, to investigate the relation between the different language proficiency and memory ability measurements and gray matter structure, whole-brain regression analyses were performed with the scores from the TROG-2, Syntest, PPVT-4, digit span, nonword repetition task, and word pairs test as a predictor of gray matter volume and cortical thickness. Again, TIV was a nuisance covariate in these regression analyses. The clusters resulting from these analyses, where gray matter structure was related to one of the language proficiency or memory ability measures, were then overlapped with the fMRI clusters.

Lastly, Menks et al. (2024) performed a breakpoint analysis on the LOESS curve, which revealed a behavioral turning point in the LOESS curve at 15.4 yr of age. This means that after the age of 15.4 yr, age had no significant effect on grammar learning success anymore. Based on this turning point, Menks et al. (2024) split the sample into a young and older group and conducted a group comparison to further explore age differences in grammatical processing before and after this turning point. Similarly, we aimed to investigate whether the relationship between brain structure and GJT performance differed between these young (8 to 15.4 yr, $n = 49$) and older (15.4 to 25.9 yr, $n = 110$) groups. Therefore, regression analyses for both groups were performed for gray matter volume and cortical thickness with the mean GJT score as a predictor and TIV as a nuisance covariate. Note that only this analysis used the raw mean GJT score as a predictor. In the other analyses described above, the predictor was always the LOESS-extracted GJT score.

Results

Behavioral results

Grammaticality judgment task

Performance on the GJT was analyzed previously by Menks et al. (2024). All participants that were included in the analyses scored above chance level in the GJT. The performance score was adjusted for each participant based on the level they reached, thus allowing us to compare the scores across the three levels of difficulty (i.e. for level 1 the score was multiplied by 0.50, for level 2 with 0.75, and for level 3 with 1). The overall grammar learning performance increased nonlinearly with age. Due to this nonlinear increase, a LOESS curve was fitted to examine the model-free relationship between age and grammar learning performance ($R^2 = 0.23$). The line of the LOESS curve followed a linear increase from age 8 until late teens after which the curve showed a slight decline followed by a modest nonsignificant increase (Fig. 1D). Additionally, a breakpoint analysis was performed to identify the age at which the effect of age on grammar learning success was no longer significant. This was reported to be at 15.4 yr of age (Menks et al. 2024).

Language proficiency and memory measurements

Mean percentage scores of the three tasks measuring L2 grammar proficiency, L1 grammar proficiency, and L2 vocabulary knowledge were as follows: TROG-2 = 93.39 (SD = 13.37, range: 25

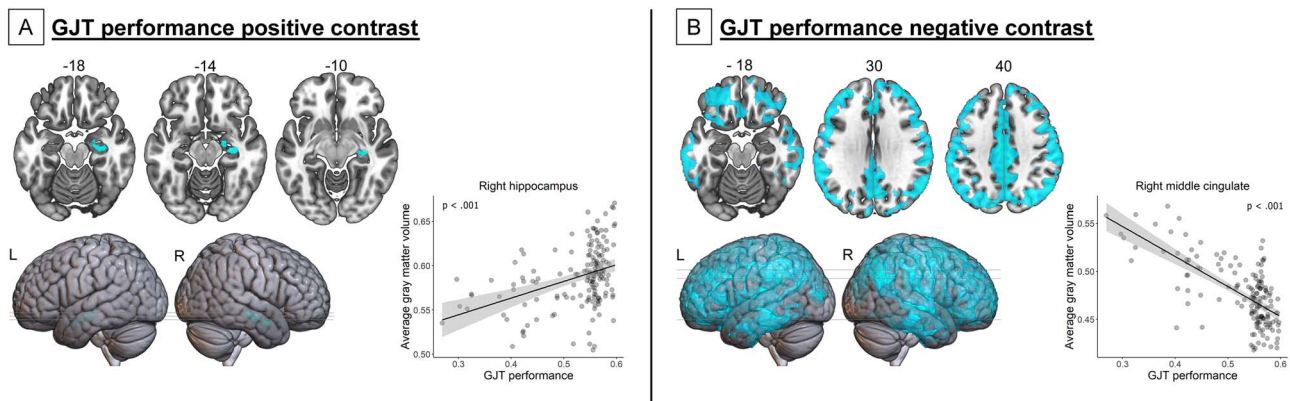


Fig. 2. Relationship between gray matter volume and LOESS-extracted GJT scores (i.e. predicted GJT performance based on age extracted from the LOESS curve). Highlighted regions show an increase (A) or decrease (B) in gray matter volume as a function of GJT performance (cluster level $P = 0.001$, $P < 0.05$ FWE-corrected). Scatterplots show the relationship between the average gray matter volume in the significant clusters and the LOESS-extracted GJT scores.

to 100), Syntest = 95.78 (SD = 3.98, range: 81.43 to 100), and PPVT-4 = 75.42 (SD = 17.49, range: 9.21 to 94.71). Mean raw scores of the memory tasks were for the nonword repetition task = 8.92 (SD = 2.06, range: 4 to 15) and for the word pairs test = 100.31 (SD = 8.22, range: 77 to 110). The mean standardized score of the digit span test was 12.81 (SD = 2.69, range: 6 to 19).

Structural MRI results

Gray matter structure and age

To explore the general developmental trajectory of brain structure in our sample, whole-brain regression analyses with age as the predictor of gray matter volume and cortical thickness were performed. This revealed a positive relationship for age and gray matter volume in the right hippocampus and the left amygdala (Fig. S2A, Table S1). A negative relationship for age and gray matter volume was found in many regions, with peaks in the paracentral lobule and the cerebellum (Fig. S2B, Table S1). Cortical thickness showed a similar pattern. Thickness in the bilateral insula had a significant positive relationship with age (Fig. S3A, Table S1). A negative relationship between cortical thickness and age was found in many brain regions, with peaks in the left postcentral gyrus, the right middle cingulum, the right parahippocampal gyrus, the right fusiform gyrus, the right lingual gyrus, and the right inferior temporal gyrus (Fig. S3B, Table S1).

Gray matter structure and age-related GJT performance

Whole-brain regression analyses for gray matter volume with age-related GJT performance as a predictor revealed a significant positive relationship in the right hippocampus (Fig. 2A, Table 1). A negative relationship between gray matter volume and age-related GJT performance was found in a large cluster, encompassing regions of the frontal, temporal, occipital, and parietal cortices, with a peak in the right middle cingulate (Fig. 2B, Table 1). The results of the cortical thickness analysis again resembled those of the volume analysis. A positive relationship with age-related GJT performance was found for cortical thickness in the bilateral insula (Table 1). A negative correlation was found in many temporal, frontal, and parietal brain regions (Fig. S3C). A detailed list of these brain regions is shown in Table 1.

Gray matter structure and GJT performance in the older and young group

In order to further explore the relationship between gray matter structure and GJT performance, the sample was split at the breakpoint of 15.4 yr into a younger and an older group. Whole-brain regressions were performed with the mean accuracy score of the GJT as a predictor of gray matter volume and cortical thickness for each group. This revealed a positive relationship between gray matter volume and GJT performance in the older group in the left middle frontal gyrus (Fig. S4, Table S2). No significant results were found for cortical thickness in the older group. Furthermore, no significant results were found for gray matter volume or cortical thickness in the young group.

Gray matter structure and language proficiency

Whole-brain regression analyses were performed to examine the relationship between gray matter structure and L1 and L2 language proficiency. The scores on the TROG-2, Syntest, and PPVT-4 were used as a predictor of gray matter volume and cortical thickness.

TROG-2

Gray matter volume was negatively related to performance on the TROG-2 in multiple brain regions. These clusters included the left superior temporal gyrus, the left supplementary motor area (SMA), the bilateral superior and middle frontal gyrus, the left middle temporal gyrus, and the right insula (Fig. 3B, Table S3). No significant positive relationship was found. Cortical thickness was positively associated with scores on the TROG-2 in the bilateral insula. A negative relationship was found in many parietal, frontal, and temporal regions (Fig. S3D, Table S3).

Syntest

Gray matter volume was negatively related to the Syntest scores in the bilateral angular gyrus, the left SMA, the right middle frontal gyrus, the left posterior cingulate, and the left superior temporal pole (Fig. 3C, Table S4). As for cortical thickness, a negative relationship was found in regions including the right supra-marginal gyrus, the bilateral superior frontal gyrus, and the right middle frontal gyrus (Fig. S3E, Table S4). No significant positive associations were found between gray matter volume or cortical thickness and the Syntest scores.

Table 1. Whole-brain regression analyses for gray matter volume and thickness with predicted grammar performance based on age (extracted from the LOESS curve) as a predictor revealed significant peak clusters for both positive (pos) and negative (neg) contrasts.

#	Region	L/R	vol	Local maxima			T	P-value
				x	y	z		
Volume								
<i>pos</i>								
1	Hippocampus, thalamus, parahippocampal gyrus	L/R	1171	31.5	-18	-15	5.29	0.000
<i>neg</i>								
1	Middle cingulate/occipital, supramarginal gyrus	L/R	179,148	3	-33	46.5	8.90	0.000
Thickness								
<i>pos</i>								
1	Insula	L	889	-30	-21	9	6.86	0.000
2	Rolandic operculum, insula	R	1074	42	-27	15	5.07	0.000
<i>neg</i>								
1	Middle cingulum, precuneus	L	9903	-8	-38	43	8.25	0.000
2	Middle cingulum, inferior parietal gyrus, precuneus	R	8987	9	-39	45	8.16	0.000
3	Superior/middle frontal gyrus (orbital)	R	2321	21	28	39.5	5.71	0.000
4	Fusiform gyrus	L	173	-28	-28	-29	5.32	0.001
5	Cerebellum, fusiform gyrus	R	311	17	-64	-20	5.17	0.000
6	Anterior cingulate, superior frontal gyrus (orbital)	L/R	399	-1	42	-2.5	5.05	0.000
7	Middle temporal gyrus	L	89	-56	-47.5	2	4.97	0.049
8	Superior/middle temporal gyrus	R	448	61	-5	-3	4.83	0.000
9	Superior frontal gyrus (orbital/medial)	L	218	-6	24	51	4.77	0.000
10	Middle temporal/superior frontal gyrus (orbital)	R	319	36	52	-0.5	4.68	0.000
11	Middle temporal gyrus	R	170	57	-49	-2	4.34	0.001
12	Supramarginal gyrus	L	104	-49.5	-39.5	23	4.33	0.025
13	Superior frontal gyrus, anterior cingulate, precentral	L	180	-17	-9	56	4.13	0.000
14	Superior/middle temporal gyrus	L	169	-56	-45	13	4.10	0.001

Note. All x-, y-, z-coordinates represent local maxima in MNI space. Cluster threshold $P < 0.001$; $P < 0.05$ FWE-corrected.

PPVT-4

Gray matter volume was positively related to PPVT-4 scores in the right hippocampus and the left amygdala (Fig. S5B, Table S5). A negative relationship between PPVT-4 performance and gray matter volume was found in a large cluster encompassing the left superior temporal gyrus and the left inferior parietal gyrus (Fig. 3D, Table S5). As for cortical thickness, PPVT-4 scores were positively related to cortical thickness in the bilateral insula (Table S5). Furthermore, a negative correlation was found for cortical thickness and scores on the PPVT-4 in frontal, temporal, parietal, and occipital areas, such as the bilateral superior frontal gyrus, the bilateral superior and middle temporal gyrus, and left middle occipital gyrus (Fig. S3F, Table S5).

Gray matter structure and memory abilities

Finally, the relationship between gray matter structure and memory abilities was examined. No significant relationships were found between gray matter volume and cortical thickness and performance on the nonword repetition task and word pairs test. Only the SMA showed a significant positive relationship between gray matter volume and performance on the digit span test (Fig. S5A, Table S6).

Overlap between structural and functional MRI results

Our second aim was to compare the structural results reported above with the fMRI clusters reported previously (Menks et al. 2024). Therefore, a mask of the fMRI clusters was overlaid with the gray matter volume clusters that we found to be negatively related to age-related GJT performance. Overlap was found in the right superior parietal gyrus, the left inferior parietal gyrus, the left precentral gyrus, and the left inferior temporal gyrus (Fig. 3E,

Table 2). Finally, the fMRI clusters were also overlaid with clusters in which gray matter volume was negatively related to L1 and L2 language proficiency. For the TROG-2 and PPVT-4, we again found overlap in the left inferior parietal gyrus and the right superior parietal gyrus (Fig. 3F–H, Table 2). For the Syntest, however, only a small region in the right angular gyrus overlapped (Fig. 3G, Table 2). In sum, the overlap with the fMRI clusters showed that in the left inferior parietal gyrus and the right superior parietal gyrus, functional activity was positively related to age-related GJT performance while at the same time gray matter volume in these same regions was negatively related to age-related GJT performance and L2 language proficiency measures.

Discussion

In this study, we investigated how developmental differences in gray matter structure relate to novel grammar learning performance in participants aged between 8 and 25 yr. Furthermore, we examined how these structural results relate to the developmental differences in brain activity during novel grammar learning reported recently by Menks et al. (2024). First, we found that GJT performance was negatively related to gray matter volume and cortical thickness in a large cluster encompassing many frontal, temporal, and parietal regions throughout the brain. A positive relationship between GJT performance and gray matter volume was only found in the right hippocampus. Next, we examined which of the brain regions where gray matter structure was negatively related to GJT performance overlapped with regions that showed age-related functional activation during the GJT (Menks et al. 2024). We found overlap in the right superior parietal gyrus, the left inferior parietal gyrus, the left precentral gyrus, and the left inferior temporal gyrus. Thus, in these regions not only brain activation but also gray matter structure was related

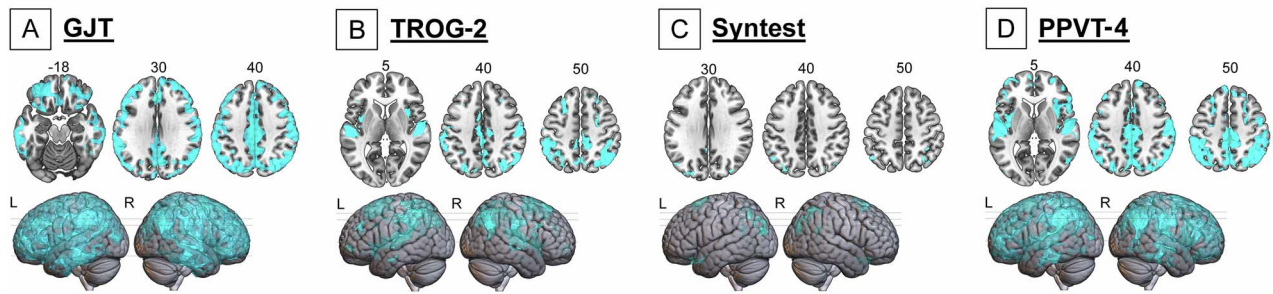
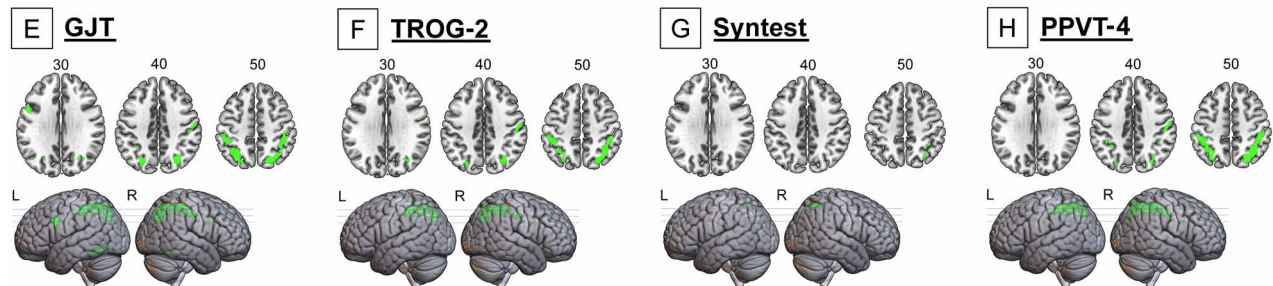
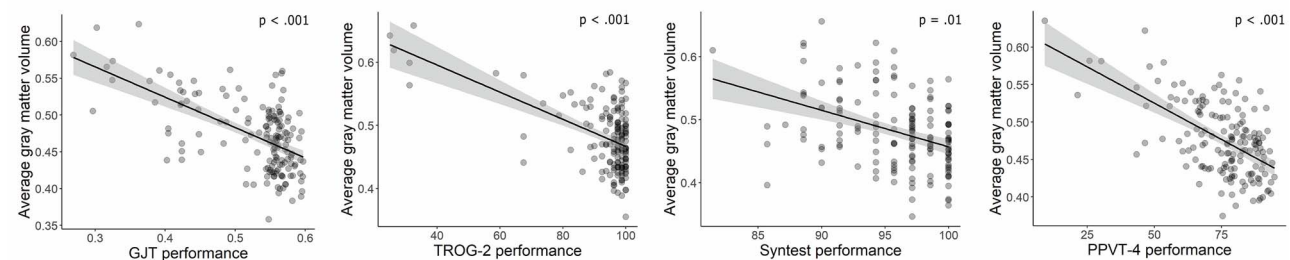
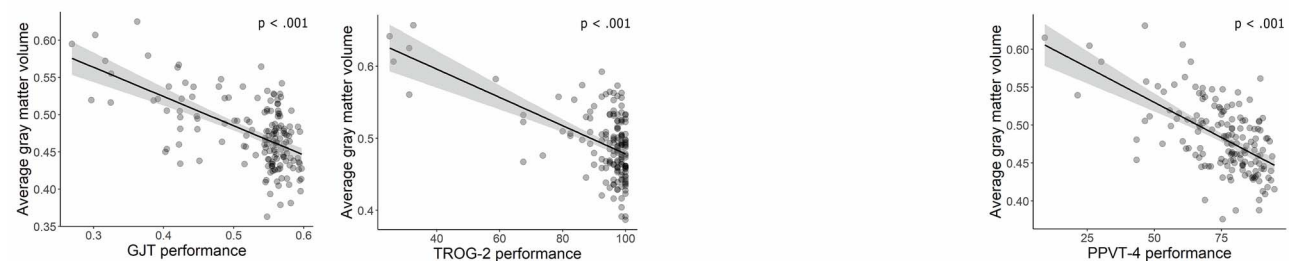
Gray matter volume negative contrast**fMRI activation overlap****Right superior parietal lobe****Left inferior parietal lobe**

Fig. 3. Relationship between gray matter volume and LOESS-extracted GJT scores (A), TROG-2 percentage scores (B), Syntest percentage scores (C), and PPVT-4 percentage scores (D). Highlighted regions show a decrease in gray matter volume as a function of test performance (cluster level $P = 0.001$, $P < 0.05$ FWE-corrected). Respective overlap with the fMRI clusters is shown in (E)–(H). Scatterplots show the relationship between the average gray matter volume in the overlapping left inferior parietal and right superior parietal clusters and test performance.

to GJT performance. Furthermore, additional analyses revealed that gray matter structure in the same regions in the left inferior parietal gyrus and the right superior parietal gyrus was related to L2 English grammar proficiency and L2 English vocabulary knowledge, but not with memory abilities. This might suggest that the left inferior parietal gyrus and the right superior parietal gyrus play an important role in second language learning. Finally, we found only minimal overlap in these regions with L1 grammar proficiency, but since performance on the Syntest was near ceiling with a restricted range, we should be cautious with interpreting these results, and we will therefore not discuss these data further.

Our first main finding was that gray matter structure was negatively related to GJT performance in many clusters throughout the brain. This was not surprising, given that both GJT performance and gray matter structure were related to age; previous analyses already showed that age was a strong predictor of GJT performance until 15.4 yr (Menks et al. 2024). Furthermore, we showed in the current study that gray matter volume and cortical thickness were negatively related to age in many clusters throughout the brain. This was expected, given that brain maturation has been related to gray matter reduction throughout the neocortex, which can already be observed at 8 yr of age or even earlier (Sowell et al. 2004a; Tamnes et al. 2010; Muftuler et al. 2011; Koolschijn and

Table 2. Overlap with fMRI clusters.

#	Region	L/R	vol	Local maxima			t	P-value
				x	y	z		
GJT								
1	Superior parietal gyrus, middle occipital gyrus	R	4174	27	-73.5	48	7.49	0.000
2	Inferior parietal gyrus, precuneus	L	3898	-46.5	-40.5	54	7.36	0.000
3	Precentral gyrus, inferior frontal gyrus (opercular), middle frontal gyrus	L	640	-51	0	28.5	6.34	0.011
4	Inferior temporal gyrus, fusiform gyrus	L	610	-48	-48	-22.5	5.08	0.14
TROG-2								
1	Inferior/superior parietal gyrus	L	2621	-48	-39	52.5	7.22	0.000
2	Superior parietal gyrus, postcentral gyrus	R	3408	31.5	-67.5	51	6.28	0.000
Syntest								
1	Angular gyrus, superior/inferior parietal gyrus	R	666	38	-66	52	4.63	0.012
PPVT-4								
1	Inferior parietal gyrus, superior parietal gyrus	L	3207	-51	-40	51	8.19	0.000
2	Superior parietal gyrus, middle occipital gyrus	R	4088	40	-52	56	7.39	0.000

Note. Clusters are shown in which GJT, TROG-2, Syntest, and PPVT-4 scores were negatively related to gray matter volume. In all these clusters, GJT performance was positively related to fMRI activation during the GJT. All x-, y-, z-coordinates represent local maxima in MNI space. Cluster threshold $P < 0.001$; $P < 0.05$ FWE-corrected. The number of voxels in each overlapping region is reported under vol.

Crone 2013; Wierenga et al. 2014; Fjell et al. 2015; Mills et al. 2016; Vidal-Pineiro et al. 2020; Bethlehem et al. 2022). Therefore, it was expected that most brain regions would follow this general developmental trend of gray matter decrease and that age-related GJT performance would be negatively related to gray matter structure in many regions.

Contrary to most brain regions, we found that hippocampal volume was positively related to GJT performance and L2 vocabulary knowledge. Although this suggests that the hippocampus follows the opposite trend of what we see in most brain regions, this finding is in line with previous studies that reported developmental increases in the volume of hippocampal (sub)regions (Giedd et al. 1996; Gogtay et al. 2006; Østby et al. 2009; Deniz Can et al. 2013). Moreover, increases in hippocampal volume have been linked previously to language learning across different ages. Deniz Can et al. (2013) observed in infants a strong positive relationship between gray matter concentration in the right hippocampus at 7 mo and expressive language ability at 12 mo. Mårtensson et al. (2012) observed in young adults an increase in hippocampal volume after 3 mo of intense foreign language training. They also found that volume of the right hippocampus was positively related to proficiency in the foreign language. From these findings, they concluded that plasticity of the hippocampus seems to be a key component for successful language learning. Even in older adults aged between 48 and 84 yr, Voits et al. (2022) found a positive relationship between bilingual experience and hippocampal volume. These studies suggest an important link between hippocampal volume and language learning throughout development and even into old age.

The relationship between hippocampal volume and language learning might reflect the involvement of declarative memory processes, which is supported by the hippocampus and other medial temporal lobe structures. According to Ullman's (2004) Declarative/Procedural model, second language learning of vocabulary and grammar both depend strongly on declarative memory. In particular, the theory predicts that adults rely more on declarative memory in the initial stages of grammar learning. Thus, an increase in hippocampal volume might be related to better declarative memory, which in turn supports second language learning. The fact that the initial stages of L2 grammar learning seem to rely heavily on declarative memory (Ullman

2004; Hamrick et al. 2018) might also explain why we did not find a relation between GJT performance and structures supporting procedural memory such as the basal ganglia. A relation with declarative memory processes is further supported by our findings for cortical thickness. We found an association between GJT performance, L2 grammar, and L2 vocabulary performance on the one hand and cortical thickness in the bilateral insula on the other hand, a structure that has also been related to declarative memory (Chen et al. 2009; Bermudez-Rattoni 2014). Surprisingly, in the current study, performance on our direct measure for declarative memory (i.e. the word pairs test) was not related to hippocampal structure, nor gray matter structure in any other brain region. One explanation could be that our declarative memory test was not sensitive enough since it was originally designed to test neuro-atypical patient populations. Indeed, many participants scored near or at ceiling level, especially in the delayed recall part of the word pairs test. Another possibility is that the relationship between hippocampal volume and GJT performance does not reflect the involvement of declarative memory processes, but involvement in another aspect of language learning. It has been proposed, for example, that the hippocampus also plays an important role in statistical learning (Schlichting et al. 2017; Finn et al. 2019).

Besides the right hippocampus, we also found that gray matter volume in the left MFG was positively related to GJT performance in the older group (15.4 to 25.9 yr). Since GJT performance was not related to age in the older group, the relationship between MFG volume and GJT performance seems to be driven by individual differences in language learning ability in this group. Previous studies have also linked the MFG to second language learning. Functional MRI studies have found task-based activation in this region during artificial grammar learning (Seeger et al. 2000; Skosnik et al. 2002; Morgan-Short et al. 2015; Kepinska et al. 2017, 2018) and L2 semantic and syntactic processing (Luke et al. 2002; Wartenburger et al. 2003). Furthermore, Mårtensson et al. (2012) found that gray matter increases in the MFG were related to foreign language training. Since the MFG has been associated with aspects of executive control, attention, and working memory (Miller 2000; Andersson et al. 2009), it has been proposed that the MFG might be involved in extralinguistic processing during language learning (Kepinska et al. 2017). In the young group (8

to 15.4 yr), we did not find a relationship between MFG structure and GJT performance, which might be explained by the fact that frontal structures, and the executive functions they support, are still developing in this group (Fuster 2002).

The second aim of the study was to examine whether our structural results (i.e. regions where we found a relationship between gray matter structure and age-related GJT performance) would overlap with the fMRI clusters recently reported by our group. We indeed observed overlap in the left inferior parietal gyrus, the right superior parietal gyrus, the left precentral gyrus, and the left inferior temporal gyrus. Thus, in these regions both structural and functional developmental differences are related to age-related GJT performance. We additionally examined whether gray matter structure in these overlapping regions was related to L2 grammar proficiency, L2 vocabulary knowledge, and memory ability. In none of these regions was gray matter structure related to memory abilities (as measured by the digit span test, nonword repetition task, and word pairs task). Interestingly, we found that L2 grammar proficiency and L2 vocabulary knowledge were related to gray matter structure in the left inferior parietal region and the right superior parietal region, which overlapped to a great extent with the fMRI clusters. This finding seems to suggest that these parietal regions are not related to general memory abilities, but to second language learning more specifically. It is also interesting to note that we found this overlap despite the fact that our language tasks tested different languages (i.e. Icelandic and English).

The involvement of the parietal cortex in language learning is well established. Already in the early 1900s, Pötzl (1925) called the lower left parietal region a “language talent area” since lesions in this area disrupted the ability to switch between languages. Similarly, Della Rosa et al. (2013) described this area as a location for multilingual talent. Mechelli et al. (2004) observed that increased gray matter density in this region was associated with early bilingualism. These findings are supported by recent neuroimaging studies that have associated activation in parietal regions with different aspects of language learning. For example, multiple studies have related parietal regions to (artificial) grammar learning (Seger et al. 2000; Skosnik et al. 2002; Tettamanti et al. 2002; Petersson et al. 2004; Forkstam et al. 2006; Finn et al. 2013; Hultén et al. 2014; Kepinska et al. 2017; Tagarelli et al. 2019; Orpella et al. 2020). Kepinska et al. (2017) reported that right parietal regions were recruited more in highly skilled learners during an artificial grammar learning task compared to moderately skilled learners, and that left angular gyrus activation was positively correlated with performance in the artificial grammar learning task. They concluded that the successful acquisition of a new grammar appears to be highly dependent on the involvement of parietal regions. Moreover, Orpella et al. (2020) established a causal link between activity in the left parietal lobe and artificial grammar learning. They found not only that activity in the left parietal lobe was positively related to statistical rule learning, but also that repetitive transcranial magnetic stimulation (rTMS) over the left parietal lobe interfered with generalizing learned rules to a new language following the same rules. In addition to artificial grammar learning, a variety of other language learning measures have been associated with the plasticity and activation of the parietal cortex, such as lexical learning and vocabulary knowledge (Cornelissen et al. 2004; Lee et al. 2007; Mestres-Missé et al. 2009; Richardson and Price 2009; Richardson et al. 2010; Grogan et al. 2012; Li et al. 2014), L2 proficiency in response to L2 training (Barbeau et al. 2017) and phonetic learning (Golestani and Zatorre 2004; Golestani and Pallier 2007). Therefore, it seems that

the parietal cortex has diverse language functions and is crucial for several aspects of language learning.

Based on the involvement of the inferior parietal lobe (IPL) in different language tasks, Shalom and Poeppel (2008) proposed a subdivision of the IPL into three subregions in which different functional aspects of language are processed. This tripartition aligns with the connectivity patterns (Caspers et al. 2011), cytoarchitectonic organization (Caspers et al. 2006), and the transmitter receptor-based organization (Caspers et al. 2013) of the IPL. Shalom and Poeppel (2008) proposed that the rostral areas of the IPL process sounds and single phonemes, the middle areas are involved in morphological operations, and the caudal areas determine the semantic content of words and sentences. The results from the current study seem to fit with this general division of the IPL; the parietal regions that we found to be associated with GJT performance seem to be located in the middle and caudal part of the inferior parietal cortex. More specifically, they appear to be ventral parts of IPL areas PF, PFm, PGa, and PGp (Caspers et al. 2006, 2011), bordering the intraparietal sulcus and the superior parietal lobe (SPL).

In order to gain a better insight into the exact functions of the left IPL and right SPL cluster, we entered the peak MNI coordinates of the observed main clusters and subclusters in Neurosynth (for the left IPL: $-46, -40, 54; -37, -61, 55; -15, -70, 60$. For the right SPL: $27, -73, 48; 31, -81, 36; 31, -67, 51$). This revealed that these (bilateral) peak coordinates have been associated with diverse language functions; many studies have reported these coordinates for tasks involving L1 and L2 semantic processing (Raboyeau et al. 2004; Tatsuno and Sakai 2005; Ketteler et al. 2008; Stein et al. 2009; Liu et al. 2010; Somers et al. 2011; van de Meerendonk et al. 2011; Buchweitz et al. 2012; Grande et al. 2012; Li et al. 2013; Liu et al. 2013; Marcotte et al. 2013; Rao et al. 2013; Segaert et al. 2013; Simard et al. 2013; Waldron and Hernandez 2013; Geranmayeh et al. 2014; Mei et al. 2014; Wang et al. 2015), as well as phonological processing (Cao et al. 2013; Koyama et al. 2013; Simard et al. 2013; Andin et al. 2015; Kim et al. 2016; Oron et al. 2016; Zhuang et al. 2016; Kim et al. 2017), listening to or generating verbs (Tatsuno and Sakai 2005; Kolndorfer et al. 2013; Waldron and Hernandez 2013; Di Cesare et al. 2017) and language switching (Ma et al. 2014; Zhang et al. 2014; Reverberi et al. 2015). More closely related to the current study is the fact that the right SPL coordinates have also been found to be involved in syntactic processes in the L1, such as processing determiner-noun gender incongruency (Heim et al. 2009), noun and verb inflection (Slioussar et al. 2014), processing ungrammatical sentences (Friederici et al. 2006; Yusa et al. 2011), and grammatical number agreement violation (Carreiras et al. 2010). The fact that we found this same region by looking at similar syntactic processes (judging case marking on Icelandic nouns) suggests that this specific region in the right SPL is not only involved in syntactic processing in the L1, but also in a newly learned language.

Especially interesting for the current study is that some studies have associated these coordinates with language learning; more specifically, they have been reported in relation to novel word learning (Mei et al. 2014; Ferreira et al. 2015) and artificial grammar learning (Opitz and Friederici 2003; Hauser et al. 2012; Finn et al. 2013; Goranskaya et al. 2016). In Finn et al. (2013), participants learned a miniature artificial language over the course of 4 d, after which they were scanned while making grammaticality judgments in the artificial language and their native language. They found that bilateral parietal regions, both very close to our peak coordinates, were activated when judging the artificial language. Furthermore, these areas were activated to a greater

extent in the artificial language than in their native language. In Goranskaya et al. (2016), participants learned an artificial grammar with an underlying phonological rule system. During scanning, there was a learning phase, in which participants had to try to extract phonological regularities from syllable sequences, and a test phase, in which they had to judge whether a syllable sequence was correct with respect to the rules they extracted. They found that participants who successfully extracted and applied the artificial grammar rules showed stronger activity in bilateral parietal regions, which were again very close to our clusters, compared to unsuccessful learners. Altogether, this literature search shows that the parietal regions found in the current study are involved in a wide variety of language functions, but more specifically that they play an important role in (artificial) grammar learning and other aspects of language learning. The exact role of these regions in second language learning remains to be elucidated.

A final result worth mentioning is that we did not find any overlap in the inferior frontal gyrus (IFG). The inferior frontal gyrus (IFG) was found to be activated during GJT performance by Menks et al. (2024), which was expected since the IFG is known as a core area in the language network (Friederici 2012; Deldar et al. 2020), playing a crucial role in syntactic processing (Friederici 2011; Tagarelli et al. 2019). Surprisingly, gray matter structure in the same IFG area was not found to be related to GJT performance. A possible explanation for this result is that the IFG might follow a different maturational trajectory than most brain regions; previous studies found that, contrary to many brain regions that undergo cortical thinning through maturation, the IFG and posterior perisylvian language regions undergo cortical thickening in childhood (Sowell et al. 2004a; Shaw et al. 2008), which has been related to the improvement of language skills (Lu et al. 2007; Nuñez et al. 2011). This regional nonlinear change in gray matter in the IFG, with a prepubertal increase followed by a postpubertal decrease, could explain why we did not find a relationship with GJT performance and therefore also no overlap.

There are some limitations in this study that could be addressed in future research. First, although we do believe that including English proficiency measures in the analyses reinforces the parietal findings, this approach is limited in that it does not account for other languages that our participants have experience with. Most participants reported in our language background questionnaire to know a third or even fourth language, and recent studies have shown that this diverse language experience can be quantified and related to brain activation and anatomy (e.g. Hervais-Adelman et al. 2018; DeLuca et al. 2019; Gullifer and Titone 2020; Kepinska et al. 2023a). Taking into account all languages reported by participants could have provided a more accurate picture of participants' multilingual experience and the associated brain structures and is a relevant consideration for future experimental designs.

In addition to considering multilingual experience, the typological distance of a participant's languages could also be an interesting consideration for future studies. In the current study, the languages we tested (i.e. Icelandic, English, and Dutch) have a small typological distance; they are all Germanic languages that share certain linguistic characteristics, such as lexical similarities. Icelandic was even specifically chosen in this study because the lexical overlap between Icelandic and Dutch made the word learning part of the study easier, so that participants could mainly focus on learning the Icelandic grammar rules. The Icelandic case

marking on the other hand—the main focus of the grammar training—does not exist in Dutch and English, which have lost most of their noun declension. Dutch and English nouns are both inflected for number, but usually not for gender or case (Dutch nouns do have gender, but there is no systematic noun inflection). For this reason, Icelandic was very suitable in the context of the current study. However, the typological distance between languages spoken by multilingual individuals has been shown to modulate brain activation. In particular, activation in the caudal part of the IPL has been found to be sensitive to the typological distance between participants' languages during language processing (Kepinska et al. 2023b). Therefore, it would be interesting to examine whether more distant languages would also overlap in the same regions that we found in the current study.

A methodological strength of the current study is that it combines both structural and functional brain data with a comparatively large developmental sample, a first for studies of novel grammar learning. Finding converging evidence from different imaging modalities increases confidence in our finding that parietal regions play an important role in second language learning. The overlap that we found might suggest that structural and functional maturation of these parietal regions follows the same developmental trajectory and that there might be a common underlying mechanism that influences both changes in brain structure and function. We cannot make any claims about whether there might be a causal link between structural and functional maturation and whether changes in gray matter might be caused by changes in activation or vice versa. Another strength is that our study investigated the effect of age in the initial stages of novel grammar learning. Interestingly, in the bilingual literature, parietal regions have been related to L2 proficiency even in individuals that have been bilingual for a long time (Chee et al. 2001; Mechelli et al. 2004; Grogan et al. 2012; Della Rosa et al. 2013; Pliatsikas et al. 2020), which suggests that these regions continue to support L2 learning over longer time periods. Future studies could address whether these parietal regions are also involved in later stages of language learning.

To conclude, by combining structural and functional data from a large developmental sample, we found several brain regions that seem to play an important role in second language learning, and novel grammar learning specifically. First, we observed that volume of the right hippocampus was positively related to age-related novel grammar learning. This might indicate that in the initial learning stage young adults rely more on this structure than children, which could be related to the observation of previous studies that, in the initial stages of grammar learning, adults seem to rely on declarative memory more than children. Second, we found that brain maturation of the left inferior parietal gyrus and the right superior parietal gyrus—as measured by increased activation during the GJT and less gray matter across development—was related to a better grammar learning performance and a higher L2 language proficiency. Thus, the earlier fMRI findings are reinforced by our gray matter findings, which shows that combining results from different imaging modalities might provide more insights into specialization of brain areas. Altogether, this current study shows that age-related differences in language learning ability is associated with structural and functional development in the parietal lobes and can serve as a foundation for future studies seeking to unravel the relationship between structural and functional development in relation to second language learning.

Acknowledgments

We thank all the participants for their time and our research assistants for a great job with the data collection: Romy Verhoeven, Carlo Rooth, and Katharina Gruber.

Author contributions

Merel E. E. Koning (Conceptualization, Investigation, Methodology, Visualization, Writing—original draft), Nina K. Wyman (Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft), Willeke M. Menks (Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Writing—review & editing), Clara Ekerdt (Conceptualization, Writing—review & editing), Guillén Fernández (Conceptualization, Funding acquisition, Project administration, Writing—review & editing), Evan Kidd (Conceptualization, Writing—review & editing), Kristin Lemhöfer (Conceptualization, Writing—review & editing), James M. McQueen (Conceptualization, Funding acquisition, Project administration, Writing—review & editing), and Gabriele Janzen (Conceptualization, Supervision, Writing—review & editing).

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Funding

This study is funded by the Netherlands Organization for Scientific Research (NWO) Gravitation grant “Language in Interaction” (grant number 024.001.006). The funder did not have a role in the design of the study and the collection, analysis, and interpretation of data.

Conflict of interest statement: None declared.

Data availability

Due to the young participants in our dataset, not all raw data could be shared. Data not available/the data that have been used are confidential.

References

- Andersson M, Ystad M, Lundervold A, Lundervold AJ. Correlations between measures of executive attention and cortical thickness of left posterior middle frontal gyrus – a dichotic listening study. *Behav Brain Funct.* 2009;5:41. <https://doi.org/10.1186/1744-9081-5-41>.
- Andin J, Fransson P, Rönnerberg J, Rudner M. Phonology and arithmetic in the language-calculation network. *Brain Lang.* 2015;143:97–105. <https://doi.org/10.1016/j.bandl.2015.02.004>.
- Ashburner J, Friston KJ. Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation. *NeuroImage.* 2011;55:954–967. <https://doi.org/10.1016/j.neuroimage.2010.12.049>.
- Asher J, Price B. The learning strategy of total physical response: some age differences. *Child Dev.* 1967;38:1219–1227. <https://doi.org/10.2307/1127119>.
- Bangen KJ, Kaup AR, Mirzakhani H, Wierenga CE, Jeste DV, Eyster LT. Compensatory brain activity during encoding among older adults with better recognition memory for face-name pairs: an integrative functional, structural, and perfusion imaging study. *J Int Neuropsychol Soc.* 2012;18:402–413. <https://doi.org/10.1017/S1355617712000197>.
- Barbeau EB, Chai XJ, Chen JK, Soles J, Berken J, Baum S, Watkins KE, Klein D. The role of the left inferior parietal lobule in second language learning: an intensive language training fMRI study. *Neuropsychologia.* 2017;98:169–176. <https://doi.org/10.1016/j.neuropsychologia.2016.10.003>.
- Bermudez-Rattoni F. The forgotten insular cortex: its role on recognition memory formation. *Neurobiol Learn Mem.* 2014;109:207–216. <https://doi.org/10.1016/j.nlm.2014.01.001>.
- Bethlehem RAI, Seidlitz J, White SR, Vogel JW, Anderson KM, Adamson C, Adler S, Alexopoulos GS, Anagnostou E, Areces-Gonzalez A, et al. Brain charts for the human lifespan. *Nature.* 2022;604:525–533. <https://doi.org/10.1038/s41586-022-04554-y>.
- Bishop DVM. TROG 2: test for reception of grammar-version 2. 2009:Edizioni Giunti OS, Florence.
- Bouman Z, Hendriks MPH, Kessels RPC, Aldenkamp AP. Bewerking en normering van de Wechsler memory scale-IV voor het gebruik in Nederland. *Tijdschrift voor Neuropsychologie.* 2012;7:99–106.
- Buchweitz A, Shinkareva SV, Mason RA, Mitchell TM, Just MA. Identifying bilingual semantic neural representations across languages. *Brain Lang.* 2012;120:282–289. <https://doi.org/10.1016/j.bandl.2011.09.003>.
- Cao F, Tao R, Liu L, Perfetti CA, Booth JR. High proficiency in a second language is characterized by greater involvement of the first language network: evidence from Chinese learners of English. *J Cogn Neurosci.* 2013;25:1649–1663. https://doi.org/10.1162/jocn_a_00414.
- Carreiras M, Carr L, Barber HA, Hernandez A. Where syntax meets math: right intraparietal sulcus activation in response to grammatical number agreement violations. *NeuroImage.* 2010;49:1741–1749. <https://doi.org/10.1016/j.neuroimage.2009.09.058>.
- Caspers S, Eickhoff SB, Rick T, von Kapri A, Kuhlen T, Huang R, Shah NJ, Zilles K. Probabilistic fibre tract analysis of cytoarchitecturally defined human inferior parietal lobule areas reveals similarities to macaques. *NeuroImage.* 2011;58:362–380. <https://doi.org/10.1016/j.neuroimage.2011.06.027>.
- Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K. The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. *NeuroImage.* 2006;33:430–448. <https://doi.org/10.1016/j.neuroimage.2006.06.054>.
- Caspers S, Schleicher A, Bacha-Trams M, Palomero-Gallagher N, Amunts K, Zilles K. Organization of the human inferior parietal lobule based on receptor architectonics. *Cereb Cortex.* 2013;23:615–628. <https://doi.org/10.1093/cercor/bhs048>.
- Chavan CF, Mouthon M, Draganski B, van der Zwaag W, Spierer L. Differential patterns of functional and structural plasticity within and between inferior frontal gyri support training-induced improvements in inhibitory control proficiency. *Hum Brain Mapp.* 2015;36:2527–2543. <https://doi.org/10.1002/hbm.22789>.
- Chee MW, Hon N, Lee HL, Soon CS. Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage.* 2001;13:1155–1163. <https://doi.org/10.1006/nimg.2001.0781>.
- Chen S, Li L, Xu B, Liu J. Insular cortex involvement in declarative memory deficits in patients with post-traumatic stress disorder. *BMC Psychiatry.* 2009;9:39. <https://doi.org/10.1186/1471-244X-9-39>.
- Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc.* 1988;83:596–610. <https://doi.org/10.1080/01621459.1988.10478639>.
- Cornelissen K, Laine M, Renvall K, Saarinen T, Martin N, Salmelin R. Learning new names for new objects: cortical effects as

- measured by magnetoencephalography. *Brain Lang.* 2004;89:617–622. <https://doi.org/10.1016/j.bandl.2003.12.007>.
- de Vos JF, Schriefers H, ten Bosch L, Lemhöfer K. Interactive L2 vocabulary acquisition in a lab-based immersion setting. *Lang Cogn Neurosci.* 2019;34:916–935. <https://doi.org/10.1080/23273798.2019.1599127>.
- Deldar Z, Gevers-Montoro C, Khatibi A, Ghazi-Saidi L. The interaction between language and working memory: a systematic review of fMRI studies in the past two decades. *AIMS Neurosci.* 2020;8:1–32. <https://doi.org/10.3934/Neuroscience.2021001>.
- Della Rosa PA, Videsott G, Borsa VM, Canini M, Weekes BS, Franceschini R, Abutalebi J. A neural interactive location for multilingual talent. *Cortex.* 2013;49:605–608. <https://doi.org/10.1016/j.cortex.2012.12.001>.
- DeLuca V, Rothman J, Bialystok E, Pliatsikas C. Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proc Natl Acad Sci USA.* 2019;116:7565–7574. <https://doi.org/10.1073/pnas.1811513116>.
- Deniz Can D, Richards T, Kuhl PK. Early gray-matter and white-matter concentration in infancy predict later language skills: a whole brain voxel-based morphometry study. *Brain Lang.* 2013;124:34–44. <https://doi.org/10.1016/j.bandl.2012.10.007>.
- Di Cesare G, Errante A, Marchi M, Cuccio V. Language for action: motor resonance during the processing of human and robotic voices. *Brain Cogn.* 2017;118:118–127. <https://doi.org/10.1016/j.bandc.2017.08.001>.
- Dunn LM, Dunn DM. *Peabody picture vocabulary test*, 2007:4th edn. Pearson, Bloomington (MN).
- Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods.* 2009;41:1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>.
- Fengler A, Meyer L, Friederici AD. Brain structural correlates of complex sentence comprehension in children. *Dev Cogn Neurosci.* 2015;15:48–57. <https://doi.org/10.1016/j.dcn.2015.09.004>.
- Ferman S, Karni A. No childhood advantage in the acquisition of skill in using an artificial language rule. *PLoS One.* 2010;5:e13648. <https://doi.org/10.1371/journal.pone.0013648>.
- Ferreira RA, Göbel SM, Hymers M, Ellis AW. The neural correlates of semantic richness: evidence from an fMRI study of word learning. *Brain Lang.* 2015;143:69–80. <https://doi.org/10.1016/j.bandl.2015.02.005>.
- Finn AS, Hudson Kam CL, Ettliger M, Vytlačil J, D'Esposito M. Learning language with the wrong neural scaffolding: the cost of neural commitment to sounds. *Front Syst Neurosci.* 2013;7:85. <https://doi.org/10.3389/fnsys.2013.00085>.
- Finn AS, Kharitonova M, Holtby N, Sheridan MA. Prefrontal and hippocampal structure predict statistical learning ability in early childhood. *J Cogn Neurosci.* 2019;31:126–137. https://doi.org/10.1162/jocn_a_01342.
- Fjell AM, Grydeland H, Krogsrud SK, Amlien I, Rohani DA, Ferschmann L, Storsve AB, Tamnes CK, Sala-Llonch R, Due-Tønnessen P, et al. Development and aging of cortical thickness correspond to genetic organization patterns. *Proc Natl Acad Sci USA.* 2015;112:15462–15467. <https://doi.org/10.1073/pnas.1508831112>.
- Flege JE. Second language speech learning: theory, findings, and problems. In: *Speech perception and linguistic experience: issues in cross-language research* Strange W (ed). York Press, Baltimore; 1995, pp 233–277.
- Forkstam C, Hagoort P, Fernández G, Ingvar M, Petersson KM. Neural correlates of artificial syntactic structure classification. *NeuroImage.* 2006;32:956–967. <https://doi.org/10.1016/j.neuroimage.2006.03.057>.
- Friederici AD. The brain basis of language processing: from structure to function. *Physiol Rev.* 2011;91:1357–1392. <https://doi.org/10.1152/physrev.00006.201>.
- Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn Sci.* 2012;16:262–268. <https://doi.org/10.1016/j.tics.2012.04.001>.
- Friederici AD, Fiebach CJ, Schlesewsky M, Bornkessel ID, von Cramon DY. Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb Cortex.* 2006;16:1709–1717. <https://doi.org/10.1093/cercor/bhj106>.
- Fuster JM. Frontal lobe and cognitive development. *J Neurocytol.* 2002;31:373–385. <https://doi.org/10.1023/a:1024190429920>.
- Geranmayeh F, Wise RJ, Mehta A, Leech R. Overlapping networks engaged during spoken language production and its cognitive control. *J Neurosci.* 2014;34:8728–8740. <https://doi.org/10.1523/JNEUROSCI.0428-14.2014>.
- Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, Paus T, Evans AC, Rapoport JL. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci.* 1999;2:861–863. <https://doi.org/10.1038/13158>.
- Giedd JN, Vaituzis AC, Hamburger SD, Lange N, Rajapakse JC, Kaysen D, Vauss YC, Rapoport JL. Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: ages 4–18 years. *J Comp Neurol.* 1996;366:223–230. [https://doi.org/10.1002/\(SICI\)1096-9861\(19960304\)366:2<#x003C;223::AID-CNE3>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1096-9861(19960304)366:2<#x003C;223::AID-CNE3>3.0.CO;2-7).
- Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TF 3rd, Herman DH, Clasen LS, Toga AW, et al. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci USA.* 2004;101:8174–8179. <https://doi.org/10.1073/pnas.0402680101>.
- Gogtay N, Nugent TF 3rd, Herman DH, Ordóñez A, Greenstein D, Hayashi KM, Clasen L, Toga AW, Giedd JN, Rapoport JL, Thompson PM. Dynamic mapping of normal human hippocampal development. *Hippocampus.* 2006;16:664–672. <https://doi.org/10.1002/hipo.20193>.
- Golestani N, Pallier C. Anatomical correlates of foreign speech sound production. *Cereb Cortex.* 2007;17:929–934. <https://doi.org/10.1093/cercor/bhl003>.
- Golestani N, Zatorre RJ. Learning new sounds of speech: reallocation of neural substrates. *NeuroImage.* 2004;21:494–506. <https://doi.org/10.1016/j.neuroimage.2003.09.071>.
- Goranskaya D, Kreitewolf J, Mueller JL, Friederici AD, Hartwigsen G. Fronto-parietal contributions to phonological processes in successful artificial grammar learning. *Front Hum Neurosci.* 2016;10:551. <https://doi.org/10.3389/fnhum.2016.00551>.
- Grande M, Meffert E, Schoenberger E, Jung S, Frauenrath T, Huber W, Hussmann K, Moormann M, Heim S. From a concept to a word in a syntactically complete sentence: an fMRI study on spontaneous language production in an overt picture description task. *NeuroImage.* 2012;61:702–714. <https://doi.org/10.1016/j.neuroimage.2012.03.087>.
- Grogan A, Parker Jones O, Ali N, Crinion J, Orabona S, Mechias ML, Ramsden S, Green DW, Price CJ. Structural correlates for lexical efficiency and number of languages in non-native speakers of English. *Neuropsychologia.* 2012;50:1347–1352. <https://doi.org/10.1016/j.neuropsychologia.2012.02.019>.
- Gullifer JW, Titone D. Characterizing the social diversity of bilingualism using language entropy. *Biling Lang Cogn.* 2020;23:283–294. <https://doi.org/10.1017/S1366728919000026>.

- Gur RC, Butler ER, Moore TM, Rosen AFG, Ruparel K, Satterthwaite TD, Roalf DR, Gennatas ED, Bilker WB, Shinohara RT, et al. Structural and functional brain parameters related to cognitive performance across development: replication and extension of the parieto-frontal integration theory in a single sample. *Cereb Cortex*. 2021;31:1444–1463. <https://doi.org/10.1093/cercor/bhaa282>.
- Hamrick P, Lum JAG, Ullman MT. Child first language and adult second language are both tied to general-purpose learning systems. *Proc Natl Acad Sci*. 2018;115:1487–1492. <https://doi.org/10.1073/pnas.1713975115>.
- Hartshorne JK, Tenenbaum JB, Pinker S. A critical period for second language acquisition: evidence from 2/3 million English speakers. *Cognition*. 2018;177:263–277. <https://doi.org/10.1016/j.cognition.2018.04.007>.
- Hashizume H, Taki Y, Sassa Y, Thyreau B, Asano M, Asano K, Takeuchi H, Nouchi R, Kotozaki Y, Jeong H, et al. Developmental changes in brain activation involved in the production of novel speech sounds in children. *Hum Brain Mapp*. 2014;35:4079–4089. <https://doi.org/10.1002/hbm.22460>.
- Hauser MF, Hofmann J, Opitz B. Rule and similarity in grammar: their interplay and individual differences in the brain. *NeuroImage*. 2012;60:2019–2026. <https://doi.org/10.1016/j.neuroimage.2012.02.016>.
- Heim S, Friederici AD, Schiller NO, Rüschemeyer SA, Amunts K. The determiner congruency effect in language production investigated with functional MRI. *Hum Brain Mapp*. 2009;30:928–940. <https://doi.org/10.1002/hbm.20556>.
- Hervais-Adelman A, Egorova N, Golestani N. Beyond bilingualism: multilingual experience correlates with caudate volume. *Brain Struct Funct*. 2018;223:3495–3502. <https://doi.org/10.1007/s00429-018-1695-0>.
- Hosoda C, Tanaka K, Nariai T, Honda M, Hanakawa T. Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *J Neurosci*. 2013;33:13663–13672. <https://doi.org/10.1523/JNEUROSCI.0410-13.2013>.
- Hu S, Chao HH, Zhang S, Ide JS, Li CS. Changes in cerebral morphology and amplitude of low-frequency fluctuations of BOLD signals during healthy aging: correlation with inhibitory control. *Brain Struct Funct*. 2014;219:983–994. <https://doi.org/10.1007/s00429-013-0548-0>.
- Hultén A, Karvonen L, Laine M, Salmelin R. Producing speech with a newly learned morphosyntax and vocabulary: an magnetoencephalography study. *J Cogn Neurosci*. 2014;26:1721–1735. https://doi.org/10.1162/jocn_a_00558.
- Janssen N, de Swart PJF, Roelofs APA, Kessels RPC, Piai V. Syntest: a Dutch test for the comprehension of (complex) syntactic structures. Unpublished.
- Johnson JS, Newport EL. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn Psychol*. 1989;21:60–99. [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0).
- Kepinska O, Dalboni da Rocha JL, Tuerk C, Hervais-Adelman A, Bouhali F, Green D, Price CJ, Golestani N. Auditory cortex anatomy reflects multilingual phonological experience. *elife*. 2023a;12:RP90269. <https://doi.org/10.7554/elife.90269.1>.
- Kepinska O, Caballero J, Oliver M, Marks RA, Haft SL, Zekelman L, Kovelman I, Uchikoshi Y, Hoeft F. Language combinations of multilinguals are reflected in their first-language knowledge and processing. *Sci Rep*. 2023b;13:1947. <https://doi.org/10.1038/s41598-023-27952-2>.
- Kepinska O, de Rover M, Caspers J, Schiller NO. On neural correlates of individual differences in novel grammar learning: an fMRI study. *Neuropsychologia*. 2017;98:156–168. <https://doi.org/10.1016/j.neuropsychologia.2016.06.014>.
- Kepinska O, de Rover M, Caspers J, Schiller NO. Connectivity of the hippocampus and Broca's area during acquisition of a novel grammar. *NeuroImage*. 2018;165:1–10. <https://doi.org/10.1016/j.neuroimage.2017.09.058>.
- Ketteler D, Kastrau F, Vohn R, Huber W. The subcortical role of language processing. High level linguistic features such as ambiguity-resolution and the human brain; an fMRI study. *NeuroImage*. 2008;39:2002–2009. <https://doi.org/10.1016/j.neuroimage.2007.10.023>.
- Kim SY, Liu L, Cao F. How does first language (L1) influence second language (L2) reading in the brain? Evidence from Korean-English and Chinese-English bilinguals. *Brain Lang*. 2017;171:1–13. <https://doi.org/10.1016/j.bandl.2017.04.003>.
- Kim SY, Qi T, Feng X, Ding G, Liu L, Cao F. How does language distance between L1 and L2 affect the L2 brain network? An fMRI study of Korean-Chinese-English trilinguals. *NeuroImage*. 2016;129:25–39. <https://doi.org/10.1016/j.neuroimage.2015.11.068>.
- Kollndorfer K, Furtner J, Krajnik J, Prayer D, Schöpf V. Attention shifts the language network reflecting paradigm presentation. *Front Hum Neurosci*. 2013;7:809. <https://doi.org/10.3389/fnhum.2013.00809>.
- Koolschijn PC, Crone EA. Sex differences and structural brain maturation from childhood to early adulthood. *Dev Cogn Neurosci*. 2013;5:106–118. <https://doi.org/10.1016/j.dcn.2013.02.003>.
- Koyama MS, Stein JF, Stoodley CJ, Hansen PC. Cerebral mechanisms for different second language writing systems. *Neuropsychologia*. 2013;51:2261–2270. <https://doi.org/10.1016/j.neuropsychologia.2013.08.002>.
- Krashen SD, Long MA, Scarcella RC. Age, rate and eventual attainment in second language acquisition. *TESOL Q*. 1979;13:573–582. <https://doi.org/10.2307/3586451>.
- Lee H, Shakeshaft C, Stewart LH, Brennan A, Glensman J, Pitcher K, Crinion J, Mechelli A, Frackowiak RSJ, et al. Anatomical traces of vocabulary acquisition in the adolescent brain. *J Neurosci*. 2007;27:1184–1189. <https://doi.org/10.1523/JNEUROSCI.4442-06.2007>.
- Lerman-Sinkoff DB, Sui J, Rachakonda S, Kandala S, Calhoun VD, Barch DM. Multimodal neural correlates of cognitive control in the human connectome project. *NeuroImage*. 2017;163:41–54. <https://doi.org/10.1016/j.neuroimage.2017.08.081>.
- Li P, Legault J, Litcofsky KA. Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex*. 2014;58:301–324. <https://doi.org/10.1016/j.cortex.2014.05.001>.
- Li Y, Yang J, Scherf KS, Li P. Two faces, two languages: an fMRI study of bilingual picture naming. *Brain Lang*. 2013;127:452–462. <https://doi.org/10.1016/j.bandl.2013.09.005>.
- Litwińczuk MC, Trujillo-Barreto N, Muhlert N, Cloutman L, Woollams A. Relating cognition to both brain structure and function: a systematic review of methods. *Brain Connect*. 2023;13:120–132. <https://doi.org/10.1089/brain.2022.0036>.
- Liu H, Hu Z, Guo T, Peng D. Speaking words in two languages with one brain: neural overlap and dissociation. *Brain Res*. 2010;1316:75–82. <https://doi.org/10.1016/j.brainres.2009.12.030>.
- Liu C, Tardif T, Wu H, Monk CS, Luo YJ, Mai X. The representation of category typicality in the frontal cortex and its cross-linguistic variations. *Brain Lang*. 2013;127:415–427. <https://doi.org/10.1016/j.bandl.2013.09.002>.
- Lu LH, Dapretto M, O'Hare ED, Kan E, McCourt ST, Thompson PM, Toga AW, Bookheimer SY, Sowell ER. Relationships between brain activation and brain structure in normally developing children.

- Cereb Cortex*. 2009;19:2595–2604. <https://doi.org/10.1093/cercor/bhp011>.
- Lu L, Leonard C, Thompson P, Kan E, Jolley J, Welcome S, Toga A, Sowell E. Normal developmental changes in inferior frontal gray matter are associated with improvement in phonological processing: a longitudinal MRI analysis. *Cereb Cortex*. 2007;17:1092–1099. <https://doi.org/10.1093/cercor/bhl019>.
- Luke KK, Liu HL, Wai YY, Wan YL, Tan LH. Functional anatomy of syntactic and semantic processing in language comprehension. *Hum Brain Mapp*. 2002;16:133–145. <https://doi.org/10.1002/hbm.10029>.
- Ma H, Hu J, Xi J, Shen W, Ge J, Geng F, Wu Y, Guo J, Yao D. Bilingual cognitive control in language switching: an fMRI study of English-Chinese late bilinguals. *PLoS One*. 2014;9:e106468. <https://doi.org/10.1371/journal.pone.0106468>.
- Marcotte K, Perlberg V, Marrelec G, Benali H, Ansaldo AI. Default-mode network functional connectivity in aphasia: therapy-induced neuroplasticity. *Brain Lang*. 2013;124:45–55. <https://doi.org/10.1016/j.bandl.2012.11.004>.
- Marques JP, Gruetter R. New developments and applications of the MP2RAGE sequence – focusing the contrast and high spatial resolution R1 mapping. *PLoS One*. 2013;8:e69294. <https://doi.org/10.1371/journal.pone.0069294>.
- Marques JP, Kober T, Krueger G, van der Zwaag W, Van de Moortele PF, Gruetter R. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage*. 2010;49:1271–1281. <https://doi.org/10.1016/j.neuroimage.2009.10.002>.
- Mårtensson J, Eriksson J, Bodammer NC, Lindgren M, Johansson M, Nyberg L, Lövdén M. Growth of language-related brain areas after foreign language learning. *NeuroImage*. 2012;63:240–244. <https://doi.org/10.1016/j.neuroimage.2012.06.043>.
- McNealy K, Mazziotta JC, Dapretto M. Age and experience shape developmental changes in the neural basis of language-related learning. *Dev Sci*. 2011;14:1261–1282. <https://doi.org/10.1111/j.1467-7687.2011.01075.x>.
- Mechelli A, Crinion JT, Noppeney U, O’Doherty J, Ashburner J, Frackowiak RS, Price CJ. Structural plasticity in the bilingual brain. *Nature*. 2004;431:757. <https://doi.org/10.1038/431757a>.
- Mei L, Xue G, Lu Z-L, Chen C, Zhang M, He Q, Wei M, Dong Q. Learning to read words in a new language shapes the neural organization of the prior languages. *Neuropsychologia*. 2014;65:156–168. <https://doi.org/10.1016/j.neuropsychologia.2014.10.019>.
- Menks WM, Ekerdt C, Janzen G, Kidd E, Lemhöfer K, Fernández G, McQueen JM. Study protocol: a comprehensive multi-method neuroimaging approach to disentangle developmental effects and individual differences in second language learning. *BMC Psychol*. 2022;10:1–17. <https://doi.org/10.1186/s40359-022-00873-x>.
- Menks WM, Ekerdt C, Lemhöfer K, Kidd E, Fernández G, McQueen JM, Janzen G. Developmental changes in brain activation during novel grammar learning in 8-25-year-olds. *Dev Cogn Neurosci*. 2024;66:101347. <https://doi.org/10.1016/j.dcn.2024.101347>.
- Mestres-Missé A, Münte TF, Rodriguez-Fornells A. Functional neuroanatomy of contextual acquisition of concrete and abstract words. *J Cogn Neurosci*. 2009;21:2154–2171. <https://doi.org/10.1162/jocn.2008.21171>.
- Miller EK. The prefrontal cortex and cognitive control. *Nat Rev Neurosci*. 2000;1:59–65. <https://doi.org/10.1038/35036228>.
- Mills KL, Goddings AL, Herting MM, Meuwese R, Blakemore SJ, Crone EA, Dahl RE, Güroğlu B, Raznahan A, Sowell ER, Tamnes CK. Structural brain development between childhood and adulthood: convergence across four longitudinal samples. *NeuroImage*. 2016;141:273–281. <https://doi.org/10.1016/j.neuroimage.2016.07.044>.
- Morgan-Short K, Deng Z, Brill-Schuetz KA, Faretta-Stutenberg M, Wong PCM, Wong FCK. A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Stud Second Lang Acquis*. 2015;37:383–419. <https://doi.org/10.1017/S0272263115000030>.
- Muftuler LT, Davis EP, Buss C, Head K, Hasso AN, Sandman CA. Cortical and subcortical changes in typically developing preadolescent children. *Brain Res*. 2011;1399:15–24. <https://doi.org/10.1016/j.brainres.2011.05.018>.
- Norbom LB, Ferschmann L, Parker N, Agartz I, Andreassen OA, Paus T, Westlye LT, Tamnes CK. New insights into the dynamic development of the cerebral cortex in childhood and adolescence: integrating macro- and microstructural MRI findings. *Prog Neurobiol*. 2021;204:102109. <https://doi.org/10.1016/j.pneurobio.2021.102109>.
- Nuñez SC, Dapretto M, Katzir T, Starr A, Bramen J, Kan E, Bookheimer S, Sowell ER. fMRI of syntactic processing in typically developing children: structural correlates in the inferior frontal gyrus. *Dev Cogn Neurosci*. 2011;1:313–323. <https://doi.org/10.1016/j.dcn.2011.02.004>.
- Opitz B, Friederici AD. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*. 2003;19:1730–1737. [https://doi.org/10.1016/s1053-8119\(03\)00170-8](https://doi.org/10.1016/s1053-8119(03)00170-8).
- Oron A, Wolak T, Zeffiro T, Szegal E. Cross-modal comparisons of stimulus specificity and commonality in phonological processing. *Brain Lang*. 2016;155-156:12–23. <https://doi.org/10.1016/j.bandl.2016.02.001>.
- Orpella J, Ripollés P, Ruzzoli M, Amengual JL, Callejas A, Martínez-Alvarez A, Soto-Faraco S, de Diego-Balaguer R. Integrating when and what information in the left parietal lobe allows language rule generalization. *PLoS Biol*. 2020;18:e3000895. <https://doi.org/10.1371/journal.pbio.3000895>.
- Østby Y, Tamnes CK, Fjell AM, Westlye LT, Due-Tønnessen P, Walhovd KB. Heterogeneity in subcortical brain development: a structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *J Neurosci*. 2009;29:11772–11782. <https://doi.org/10.1523/JNEUROSCI.1242-09.2009>.
- Persson J, Nyberg L, Lind J, Larsson A, Nilsson LG, Ingvar M, Buckner RL. Structure-function correlates of cognitive decline in aging. *Cereb Cortex*. 2006;16:907–915. <https://doi.org/10.1093/cercor/bhj036>.
- Petersson KM, Forkstam C, Ingvar M. Artificial syntactic violations activate Broca’s region. *Cogn Sci*. 2004;28:383–407. <https://doi.org/10.1016/j.cogsci.2003.12.003>.
- Pliatsikas C, Meteyard L, Verissimo J, DeLuca V, Shattuck K, Ullman MT. The effect of bilingualism on brain development from early childhood to young adulthood. *Brain Struct Funct*. 2020;225:2131–2152. <https://doi.org/10.1007/s00429-020-02115-5>.
- Pötzl O. Über die parietal bedingte Aphasie und ihren Einfluss auf das Sprechen mehrerer Sprachen. *Z Gesamt Neurol Psychiatr*. 1925;96:100–124. <https://doi.org/10.1007/BF02896516>.
- Raboyeau G, Marie N, Balduyck S, Gros H, Démonet JF, Cardebat D. Lexical learning of the English language: a PET study in healthy French subjects. *NeuroImage*. 2004;22:1808–1818. <https://doi.org/10.1016/j.neuroimage.2004.05.011>.
- Rao C, Mathur A, Singh NC. ‘Cost in transliteration’: the neurocognitive processing of Romanized writing. *Brain Lang*. 2013;124:205–212. <https://doi.org/10.1016/j.bandl.2012.12.004>.
- Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, Clasen L, Gogtay N, Giedd JN. How does your cortex grow? *J*

- Neurosci.* 2011;31:7174–7177. <https://doi.org/10.1523/JNEUROSCI.0054-11.2011>.
- Reverberi C, Kuhlen A, Abutalebi J, Greulich RS, Costa A, Seyed-Allaei S, Haynes JD. Language control in bilinguals: intention to speak vs. execution of speech. *Brain Lang.* 2015;144:1–9. <https://doi.org/10.1016/j.bandl.2015.03.004>.
- Richardson FM, Price CJ. Structural MRI studies of language function in the undamaged brain. *Brain Struct Funct.* 2009;213:511–523. <https://doi.org/10.1007/s00429-009-0211-y>.
- Richardson FM, Thomas MS, Filippi R, Harth H, Price CJ. Contrasting effects of vocabulary knowledge on temporal and parietal brain structure across lifespan. *J Cogn Neurosci.* 2010;22:943–954. <https://doi.org/10.1162/jocn.2009.21238>.
- Rivera SM, Reiss AL, Eckert MA, Menon V. Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cereb Cortex.* 2005;15:1779–1790. <https://doi.org/10.1093/cercor/bhi055>.
- Salami A, Rieckmann A, Fischer H, Bäckman L. A multivariate analysis of age-related differences in functional networks supporting conflict resolution. *NeuroImage.* 2014;86:150–163. <https://doi.org/10.1016/j.neuroimage.2013.08.002>.
- Schlichting ML, Guarino KF, Schapiro AC, Turk-Browne NB, Preston AR. Hippocampal structure predicts statistical learning and associative inference abilities during development. *J Cogn Neurosci.* 2017;29:37–51. https://doi.org/10.1162/jocn_a_01028.
- Segaert K, Kempen G, Petersson KM, Hagoort P. Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: an fMRI study. *Brain Lang.* 2013;124:174–183. <https://doi.org/10.1016/j.bandl.2012.12.003>.
- Seger CA, Prabhakaran V, Poldrack RA, Gabrieli JDE. Neural activity differs between explicit and implicit learning of artificial grammar strings: an fMRI study. *Psychobiology.* 2000;28:283–292. <https://doi.org/10.3758/BF03331987>.
- Shalom DB, Poeppel D. Functional anatomic models of language: assembling the pieces. *Neuroscientist.* 2008;14:119–127. <https://doi.org/10.1177/1073858407305726>.
- Shaw P, Kabani NJ, Lerch JP, Eckstrand K, Lenroot R, Gogtay N, Greenstein D, Clasen L, Evans A, Rapoport JL, et al. Neurodevelopmental trajectories of the human cerebral cortex. *J Neurosci.* 2008;28:3586–3594. <https://doi.org/10.1523/JNEUROSCI.5309-07.2008>.
- Simard F, Monetta L, Nagano-Saito A, Monchi O. A new lexical card-sorting task for studying fronto-striatal contribution to processing language rules. *Brain Lang.* 2013;125:295–306. <https://doi.org/10.1016/j.bandl.2011.08.002>.
- Skosnik PD, Mirza F, Gitelman DR, Parrish TB, Mesulam MM, Reber PJ. Neural correlates of artificial grammar learning. *NeuroImage.* 2002;17:1306–1314. <https://doi.org/10.1006/nimg.2002.1291>.
- Slioussar N, Kireev MV, Chernigovskaya TV, Kataeva GV, Korotkov AD, Medvedev SV. An ER-fMRI study of Russian inflectional morphology. *Brain Lang.* 2014;130:33–41. <https://doi.org/10.1016/j.bandl.2014.01.006>.
- Snow C, Hoefnagel-Hohle M. The critical period for language acquisition: evidence from second language acquisition. *Child Dev.* 1978;49:1114–1128. <https://doi.org/10.2307/1128751>.
- Somers M, Neggers SF, Diederer KM, Boks MP, Kahn RS, Sommer IE. The measurement of language lateralization with functional transcranial Doppler and functional MRI: a critical evaluation. *Front Hum Neurosci.* 2011;5:31. <https://doi.org/10.3389/fnhum.2011.00031>.
- Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW. Mapping cortical change across the human life span. *Nat Neurosci.* 2003;6:309–315. <https://doi.org/10.1038/nn1008>.
- Sowell ER, Thompson PM, Holmes CJ, Batth R, Jernigan TL, Toga AW. Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *NeuroImage.* 1999a;9:587–597. <https://doi.org/10.1006/nimg.1999.0436>.
- Sowell ER, Thompson PM, Holmes CJ, Jernigan TL, Toga AW. In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nat Neurosci.* 1999b;2:859–861. <https://doi.org/10.1038/13154>.
- Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, Toga AW. Longitudinal mapping of cortical thickness and brain growth in normal children. *J Neurosci.* 2004a;24:8223–8231. <https://doi.org/10.1523/JNEUROSCI.1798-04.2004>.
- Sowell ER, Thompson PM, Tessner KD, Toga AW. Mapping continued brain growth and gray matter density reduction in dorsal frontal cortex: inverse relationships during postadolescent brain maturation. *J Neurosci.* 2001;21:8819–8829. <https://doi.org/10.1523/JNEUROSCI.21-22-08819.2001>.
- Sowell ER, Thompson PM, Toga AW. Mapping changes in the human cortex throughout the span of life. *Neuroscientist.* 2004b;10:372–392. <https://doi.org/10.1177/1073858404263960>.
- Squeglia LM, McKenna BS, Jacobus J, Castro N, Sorg SF, Tapert SF. BOLD response to working memory not related to cortical thickness during early adolescence. *Brain Res.* 2013;1537:59–68. <https://doi.org/10.1016/j.brainres.2013.08.053>.
- Stein M, Federspiel A, Koenig T, Wirth M, Lehmann C, Wiest R, Strik W, Brandeis D, Dierks T. Reduced frontal activation with increasing 2nd language proficiency. *Neuropsychologia.* 2009;47:2712–2720. <https://doi.org/10.1016/j.neuropsychologia.2009.05.023>.
- Sun J, Chen Q, Zhang Q, Li Y, Li H, Wei D, Yang W, Qiu J. Training your brain to be more creative: brain functional and structural changes induced by divergent thinking training. *Hum Brain Mapp.* 2016;37:3375–3387. <https://doi.org/10.1002/hbm.23246>.
- Tagarelli KM, Shattuck KF, Turkeltaub PE, Ullman MT. Language learning in the adult brain: a neuroanatomical meta-analysis of lexical and grammatical learning. *NeuroImage.* 2019;193:178–200. <https://doi.org/10.1016/j.neuroimage.2019.02.061>.
- Takashima A, Bakker-Marshall I, van Hell JG, McQueen JM, Janzen G. Neural correlates of word learning in children. *Dev Cogn Neurosci.* 2019;37:100649. <https://doi.org/10.1016/j.dcn.2019.100649>.
- Tamnes CK, Østby Y, Fjell AM, Westlye LT, Due-Tønnessen P, Walhovd KB. Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb Cortex.* 2010;20:534–548. <https://doi.org/10.1093/cercor/bhp118>.
- Tatsuno Y, Sakai KL. Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *J Neurosci.* 2005;25:1637–1644. <https://doi.org/10.1523/JNEUROSCI.3978-04.2005>.
- Tavakol S, Li Q, Royer J, Vos de Wael R, Larivière S, Lowe A, Paquola C, Jefferies E, Hartley T, Bernasconi A, et al. A structure-function substrate of memory for spatial configurations in medial and lateral temporal cortices. *Cereb Cortex.* 2021;31:3213–3225. <https://doi.org/10.1093/cercor/bhab001>.
- Tettamanti M, Alkadhi H, Moro A, Perani D, Kollias S, Weniger D. Neural correlates for the acquisition of natural language syntax. *NeuroImage.* 2002;17:700–709. <https://doi.org/10.1006/nimg.2002.1201>.
- Ullman MT. Contributions of memory circuits to language: the declarative/procedural model. *Cognition.* 2004;92:231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>.
- van de Meerendonk N, Indefrey P, Chwilla DJ, Kolk HH. Monitoring in language perception: electrophysiological and hemodynamic

- responses to spelling violations. *NeuroImage*. 2011;54:2350–2363. <https://doi.org/10.1016/j.neuroimage.2010.10.022>.
- Vidal-Pineiro D, Parker N, Shin J, French L, Grydeland H, Jackowski AP, Mowinckel AM, Patel Y, Pausova Z, Salum G, et al. Cellular correlates of cortical thinning throughout the lifespan. *Sci Rep*. 2020;10:21803. <https://doi.org/10.1038/s41598-020-78471-3>.
- Voits T, Robson H, Rothman J, Pliatsikas C. The effects of bilingualism on hippocampal volume in ageing bilinguals. *Brain Struct Funct*. 2022;227:979–994. <https://doi.org/10.1007/s00429-021-02436-z>.
- Waldron EJ, Hernandez AE. The role of age of acquisition on past tense generation in Spanish-English bilinguals: an fMRI study. *Brain Lang*. 2013;125:28–37. <https://doi.org/10.1016/j.bandl.2013.01.002>.
- Walhovd KB, Fjell AM, Giedd J, Dale AM, Brown TT. Through thick and thin: a need to reconcile contradictory results on trajectories in human cortical development. *Cereb Cortex*. 2017;27:1472–1481. <https://doi.org/10.1093/cercor/bhv301>.
- Wang X, Yang J, Yang J, Mencl WE, Shu H, Zevin JD. Language differences in the brain network for reading in naturalistic story reading and lexical decision. *PLoS One*. 2015;10:e0124388. <https://doi.org/10.1371/journal.pone.0124388>.
- Wartenburger I, Heekeren HR, Abutalebi J, Cappa SF, Villringer A, Perani D. Early setting of grammatical processing in the bilingual brain. *Neuron*. 2003;37:159–170. [https://doi.org/10.1016/s0896-6273\(02\)01150-9](https://doi.org/10.1016/s0896-6273(02)01150-9).
- Wechsler D. Wechsler Adult Intelligence Scale - fourth edition (WAIS-IV). 1–3.
- Wechsler DWISC-V-NL. 2018. *Wechsler Intelligence Scale for Children – fifth edition – Nederlandstalige bewerking*. Amsterdam: Pearson; 2008.
- Wierenga LM, Langen M, Oranje B, Durston S. Unique developmental trajectories of cortical thickness and surface area. *NeuroImage*. 2014;87:120–126. <https://doi.org/10.1016/j.neuroimage.2013.11.010>.
- Wong PC, Warrier CM, Penhune VB, Roy AK, Sadehh A, Parrish TB, Zatorre RJ. Volume of left Heschl's gyrus and linguistic pitch learning. *Cereb Cortex*. 2008;18:828–836. <https://doi.org/10.1093/cercor/bhm115>.
- Yusa N, Koizumi M, Kim J, Kimura N, Uchida S, Yokoyama S, Miura N, Kawashima R, Hagiwara H. Second-language instinct and instruction effects: nature and nurture in second-language acquisition. *J Cogn Neurosci*. 2011;23:2716–2730. <https://doi.org/10.1162/jocn.2011.21607>.
- Zhang Y, Huang P, Song Z, Fang L, Shen T, Li Y, Gong Q, Xie P. In-context language control with production tasks in bilinguals: an fMRI study. *Brain Res*. 2014;1585:131–140. <https://doi.org/10.1016/j.brainres.2014.08.027>.
- Zhuang J, Johnson MA, Madden DJ, Burke DM, Diaz MT. Age-related differences in resolving semantic and phonological competition during receptive language tasks. *Neuropsychologia*. 2016;93:189–199. <https://doi.org/10.1016/j.neuropsychologia.2016.10.016>.