Grammar acquisition in preschool children is related to white matter maturation of the dorsal language network

Cheslie C. Klein[1,2*], Philipp Berger[1,2], Charlotte Grosse Wiesmann[2,3‡] & Angela D. Friederici[1‡]

[1]Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[2]Research Group Milestones of Early Cognitive Development, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[3]Cognitive Neuroscience Lab, Department of Liberal Arts and Sciences, University of Technology Nuremberg, Germany

‡ Contributed equally

***Correspondence to:**

Cheslie C. Klein

Max Planck Institute for Human Cognitive and Brain Sciences

Department of Neuropsychology

Stephanstraße 1a, 04103 Leipzig, Germany

Email: cklein@cbs.mpg.de

Abstract

In preschool years, children take important steps in grammar acquisition, which are essential to learning their native language. A central aspect is the acquisition of the morpho-syntactic rule system, which forms an intersection between words and sentences. In adults, rule-based linguistic processes are supported by the dorsal fiber pathway to BA44, the arcuate fascicle. This pathway matures relatively late in development, raising the question of whether it already supports grammar processes in the early preschool years, or whether early grammar acquisition is supported by different, earlier-maturing fiber pathways. In two independent samples of 3- to 5-year-old children (N = 90 and N = 30), we examined the association between the maturation of fiber pathways of the language network and children's noun plural assignment as an index of their morpho-syntactic abilities. This revealed consistent differences between 3-year-olds and 4- to 5-year-olds. The 4- and 5-year-olds, but not 3-year-olds, showed a relation of morpho-syntax with both the dorsal pathway to BA44, supporting syntactic processes, and the dorsal pathway to BA6, supporting phonological processes in adults. Our results suggest that, in contrast to adults, preschool-aged children rely on both dorsal fiber pathways for morpho-syntax. This difference might point to different processing strategies reflecting the transition from phonology-based statistical learning to rule-based learning in grammar acquisition.

Keywords: Language development, Grammar acquisition, Morpho-syntax, Plural assignment, White matter, Arcuate fascicle, Language network

Introduction

Grammar is at the core of human language. When acquiring their native language, young children are faced with a difficult task. Not only do they need to learn a large number of new words from their input. They also need to extract rules governing the combination of words in a sentence as well as the combination of morphological elements of a word. It is an open question how the developing brain supports these processes during language acquisition.

While words contain the meaning, grammar rules set the relation between these words in a sentence, thereby allowing the formation of complex thoughts beyond the word-level. At the word-level grammatical rules guide the combinatorial processes of morphological elements in a word. Morpho-syntax builds a bridge between the structure of the sentence and its words by modulating the grammatical features of a given word in relation to the sentence. For example, the plural form of a noun takes on a more complex structure (e.g., dog-s) than the singular noun stem (e.g., dog) by combining it with a plural morpheme (e.g., -s). Even though the combinatorial process is on the word-level, it enables the production and comprehension at the sentence-level (Harley, 2015).

During development, the ability to process rule-based regularities in auditory sequences is present in the first months of life, reflecting statistical learning of phonologically encoded regularities (Friederici et al., 2011; Marcus et al., 1999). Sentence-level syntactic rules are detected by the age of 2.5 years (Oberecker et al., 2005; Oberecker & Friederici, 2006). By 3 years of age, children can already produce and understand a variety of sentences (Kauschke, 2012; Schipke et al., 2012; Skeide et al., 2016; Strotseva-Feinschmidt et al., 2019), indicating that basic combinatorial processes seem to be in place by that age. However, milestones of complex grammar acquisition are only reached later, between the ages of 4 to 5 years and beyond (Crain & Thornton, 2012; Skeide & Friederici, 2016). Thus, the preschool years are

considered as a take-off phase towards more complex grammar (Crain & Thornton, 2012; Fox & Grodzinsky, 1998; Guasti, 2002; Kauschke, 2012; Tomasello & Brooks, 1999).

At the neural level, the language network differs considerably between the developing and the adult brain. In adults, the neural language network consists of dedicated language regions in the inferior frontal cortex and the posterior temporal cortex known as Broca's and Wernicke's areas and the dorsally and ventrally located white matter fiber pathways interconnecting these regions (Friederici & Gierhan, 2013). In adults, grammatical processes involve BA44 in the posterior part of Broca's area in the left inferior frontal gyrus in interaction with the posterior temporal lobe on the level of words (Bulut, 2022; Leminen et al., 2019; Meykadeh et al., 2023; Ullman, 2001) as well as sentences (Goucha & Friederici, 2015; Grodzinsky et al., 2021; Zaccarella et al., 2017). The dorsal fiber pathway that connects these two brain regions, referred to as arcuate fascicle, has been suggested to support complex grammar in adults (Friederici & Gierhan, 2013; Griffiths et al., 2013; Wilson et al., 2011) including morpho-syntactic processes (Rolheiser et al., 2011). The ventral pathway, via the inferior fronto-occipital fascicle (IFOF), is proposed to be involved in semantic processes (Fridriksson et al., 2016; Friederici & Gierhan, 2013; Rolheiser et al., 2011; Saur et al., 2008). The arcuate fascicle, dorsally connecting to BA44, can be distinguished from a second dorsal fiber bundle that connects the auditory cortex in the superior temporal lobe and BA6 in the premotor cortex. This latter pathway has been suggested to be involved in auditory-to-motor mapping as, for example, in speech repetition in adults (Fernández-Miranda et al., 2015; Friederici & Gierhan, 2013; Saur et al., 2008) and babbling in infants (Hickok & Poeppel, 2004).

Developmental brain studies have reported functional changes in the allocation of the language-related brain regions (Klein et al., 2023; Skeide et al., 2014) as well as maturational changes of the different fiber tracts (Skeide et al., 2016). A functional shift from involvement of the temporal cortex to the inferior frontal cortex has been observed between 3 and 4 years

of age for sentence-level grammar processes (Klein et al., 2023). This shift towards more mature brain regions might not only underly the behavioral milestones in grammar between 3 and 4 years but might also point to alternative processing strategies for basic combinatorial processes before the age of 4 years. With respect to the fiber tracts, the dorsal pathway to BA6 and the IFOF are already detectable at birth, even though both fiber pathways still mature during childhood (Brauer et al., 2013; Perani et al., 2011). In contrast, the dorsal pathway to BA44 (i.e., the arcuate fascicle) could not be reconstructed in newborns, suggesting a structurally less mature fiber pathway compared to the other two (Perani et al., 2011). This has led to the hypothesis that early language development is predominantly supported by the dorsal pathway to BA6 and the IFOF, whereas the acquisition of complex sentence-level grammar from 4 years of age might require a more matured dorsal fiber connection to BA44 (Eichner et al., 2024; Skeide et al., 2016; Skeide & Friederici, 2016). This leaves open the question of how the maturation of the core language fiber pathways supports grammar acquisition on the word-level (i.e., morpho-syntax) in preschool-aged children.

The present study aims to clarify how the use and acquisition of morpho-syntactic rules relates to the maturation of core fiber connections of the language network in preschool years. We therefore investigated the association between developmental change in the dorsal and ventral fiber pathways of the language network and children's grammar proficiency at the word-level during the critical preschool period of 3 to 5 years. For this, we analyzed preschooler's morpho-syntactic performance in relation to structural properties of the two dorsal pathways targeting BA44 and BA6, and the ventral pathway (i.e., the IFOF), as well as a control tract, the corticospinal tract. This was investigated in two independent samples of 3- and 5-year-old children (Sample 1), and 3- and 4-year-old children (Sample 2), covering the ages where the shift in cortical brain regions and behavioral milestones in morpho-syntax have been observed. We hypothesized that children's grammar abilities would be associated

with indices of white matter maturation in the dorsal fiber pathway to BA44. In addition, we expected different patterns in this association between the 3-year-old versus the 4- and 5-year-old children, with younger children still relying on the well-matured dorsal pathway to BA6 or the IFOF.

Methods

The analysis pipeline and exclusion criteria were preregistered at AsPredicted (#90877).

Participants

For this study, MRI and behavioral data from two independent samples of monolingual German-speaking preschool children were analyzed. Sample 1 (N = 90) included 3- and 5-year-old children (3-y.o.s: N = 34, mean age = 3.63, SD = 0.24, range = 3.10 to 3.99, 12 female; 5-y.o.s: N = 56, mean age = 5.48, SD = 0.47, range = 4.01 to 6.16, 27 female). Sample 2 (N = 30) included 3- and 4-year-old children (3-y.o.s: N = 13, mean age = 3.26, SD = 0.18, range = 3.07 to 3.59, 8 female; 4-y.o.s: N = 17, mean age = 4.32, SD = 0.18, range = 4.02 to 4.58, 8 female). Children had no reported history of medical, neurological, or psychiatric disorder and no hearing or vision deficit.

These samples were subsamples from larger behavioral samples of N = 221 for the 3- and 5-year-olds of Sample 1, and N = 60 for the 3- and 4-year-olds of Sample 2. The behavioral samples included children who provided data from a standardized test battery of general language development (SETK 3-5; *Sprachentwicklungstest für drei- bis fünfjährige Kinder: SETK3–5*; Grimm, 2001), independent of whether they also underwent MRI scans. Of these, N = 88 (Sample 1) and N = 9 (Sample 2) did not participate in or aborted the MRI. Further, children were excluded if they did not provide behavioral data of any relevant SETK 3-5 subtest (Sample 1: N = 1) or showed an indication for a speech development delay in the SETK 3-5 (T-value < 35; Sample 1: N = 8, Sample 2: N = 2). Children were excluded from the MRI analyses if they received a later diagnosis of ADHD or developmental dyslexia (Sample 1: N = 16) or did not provide handedness data (Sample 1: N = 4, Sample 2: N = 10). After MRI preprocessing, children were further excluded if they showed severe motion artifacts in the MRI scans (see *MRI preprocessing*; Sample 1: N = 10, Sample 2: N = 6), or for which brain extraction, estimation of gradient direction or tractography failed (see *MRI*

preprocessing; Sample 1: N = 5, Sample 2: N = 3). For the standardization of the behavioral scores, children were further excluded from the initial behavioral sample if they did not provide data of all SETK 3-5 subtests (Sample 1: N = 18). This resulted in N = 194 3- and 5-year-old children for standardization within Sample 1 and N = 58 3- and 4-year-olds within Sample 2.

Parental consent was obtained for all children and the study was approved by the Ethics Committee at the Faculty of Medicine of the University of Leipzig (number of approval; Sample 1: 320-11-26092011, Sample 2: 090/12-ff). Data of both samples have previously been analyzed separately with regard to other research questions (Sample 1: Eichner et al., 2024; Kuhl et al., 2020, 2021; Liebig et al., 2020; Qi et al., 2021; Sample 2: Berger et al., 2021; Grosse Wiesmann et al., 2020; Grosse Wiesmann, Friederici, et al., 2017; Grosse Wiesmann, Schreiber, et al., 2017; Klein et al., 2023).

Behavioral data

Morpho-syntactic ability

The SETK 3-5 is a comprehensive diagnostic tool designed to test 3- to 5-year-old children's general language development in the domains of language comprehension, production, and memory (Grimm, 2001). To test children's grammar abilities on the word-level, we selected and preregistered the morpho-syntactic word production task (SETK 3-5 subtest 'Morphologische Regelbildung'; Grimm, 2001), which determines the acquisition level of the morpho-syntactic rule system for German plural formation.

In this task children are asked to name the plural of given nouns. For this, children are first shown a picture of an object or animal that is described using the singular form of a noun (e.g., "Schau mal, hier ist ein Auto." [engl. "Look, here is a car."]; Grimm, 2001). Then, a picture with three of the same objects or animals is shown and the child is asked to produce the plural form of the given word (e.g., "Hier kommen noch mehr dazu. Hier sind drei...?"

[engl. “Here are some more. Here are three...?”]; Grimm, 2001). Children are familiarized with the task with one practice item and then tested on ten common real nouns, which follow the various dominant plural rules of the German noun system (for further description of the German plural system see *Supplementary Methods*). The 4- to 5-year-old children were presented, in addition, with eight pseudo words similar to the Wug Test from Berko (1958) (for a list of the items see *Supplementary Table 1*; Grimm, 2001).

Children’s responses for each item were rated depending on whether a plural rule was used, and which one. A score of 2 was given for the correct plural form, of 1 for the use of a different but incorrect plural rule (e.g., “Fisch-en” instead of “Fisch-e”; Grimm, 2001). A score of zero was given for an incorrect response, for example, repetition of the singular form or a plural noun with significant changes of the word stem (e.g., “Gratel-en” instead of “Gabel-n”; Grimm, 2001), or no response at all. In the present study, we used raw values of the morpho-syntactic word production task rather than normalized scores provided by the SETK 3-5, as we were interested in relating individual developmental differences to brain structure. To account for the differences in item structure among 3- and 4- to 5-year-olds, we z-transformed these raw scores within each age group. The standardization was carried out independently within the full behavioral datasets of each sample (Sample 1: N = 194, Sample 2: N = 58). The resulting morpho-syntax score was used for the brain-behavior analyses.

In addition, we preregistered a syntactic comprehension and production score from two further subtests of the SETK 3-5 to test for children’s grammar ability on the sentence-level, which we report in the Supplementary Material.

Psychometrics

Children’s non-verbal IQ was assessed using the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III; Wechsler, 2009) in Sample 1, and the Kaufman Assessment Battery for Children (Kaufman ABC; Melchers & Preuß, 2003) in Sample 2. Children’s non-verbal

IQ scores were within the typical range for the respective age: For Sample 1, the standard scores of the non-verbal scale were used (mean = 98.33, SD = 13.30), and for Sample 2, the mean scale scores were used (mean = 10.26, SD = 1.34). In both samples, children's degree of handedness was evaluated with the German version of the Edinburgh Handedness Inventory (Oldfield, 1971).

MRI data acquisition

MRI data were obtained on a 3T Siemens scanner (Siemens MRT Trio series) equipped with a 12-channel head coil for Sample 1 and a 32-channel head coil for Sample 2.

In Sample 1, diffusion-weighted MRI (dMRI) data were acquired using the optimized monopolar Stejskal-Tanner EPI sequence (Morelli et al., 2010) along the anterior-to-posterior phase encoding (PE) direction (voxel size = 1.86×1.86×1.9 mm, TR = 8,000 ms, TE = 83.0 ms, b-value = 1,000 s/mm^2, 60 directions, GRAPPA 2). A field map was acquired after completion of each 10 images. In addition, one diffusion-weighted image and a field map was acquired along the posterior-to-anterior PE direction. In Sample 2, dMRI data were acquired using the multiplexed echo planar imaging (EPI) sequence (Feinberg et al., 2010) with a spatial resolution of 1.9 mm isotropic (TR = 4,000 ms, TE = 75.4 ms, b-value = 1,000 s/mm^2, 60 directions, GRAPPA 2). After the dMRI scan, a field map was acquired as an anatomical reference.

High-resolution 3D T1-weighted MRI data were further acquired in both samples using the MP2RAGE sequence (Marques et al., 2010) (Sample 1: voxel size = 1.3 mm isotropic, TR = 5,000 ms, TE = 2.82 ms, TI_1 = 700 ms, α_1 = 4°, TI_2 = 2,500 ms, α_2 = 5°; Sample 2: voxel size = 1.2×1.0×1.0 mm, TR = 5,000 ms, TE = 3.24 ms, TI_1 = 700 ms, α_1 = 4°, TI_2 = 2,500 ms, α_2 = 5°).

A summary table of the MRI parameters applied in the two samples can be found in the Supplementary Material (see *Supplementary Table 2*).

MRI data analysis

MRI preprocessing and analyses were performed independently within each of the two samples, unless otherwise stated.

MRI preprocessing

MRI preprocessing of the dMRI data included the following steps (Eichner et al., 2024; Fan et al., 2022). Estimation of the susceptibility-induced off-resonance field was done using FSL's *topup* tool to correct for distortions (Andersson et al., 2003; Smith et al., 2004). The underlying noise distribution was then used to correct for non-Gaussian noise biases in the data (St-Jean et al., 2020). Further, data was denoised and Gibbs ringing artefacts were removed using MRtrix3's *dwidenoise* and *mrdegibbs* (Kellner et al., 2016; Tournier et al., 2019; Veraart et al., 2016). Motion correction was applied with FSL's *eddy* (Andersson & Sotiropoulos, 2016). We additionally performed signal drift correction using a second-order polynomial fit to the data without diffusion-weighting (Vos et al., 2017).

T1-weighted MRI data were preprocessed with the following steps. Background noise was removed by combining the two inversion time images (O'Brien et al., 2014). The fsl_anat pipeline was then used to reorient and register the data to the MNI space, perform brain extraction, and segment the brain tissues (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/fsl_anat). Then, the data was aligned to ACPC using the Automatic Registration Toolbox (https://www.nitrc.org/projects/art). Finally, diffusion data were aligned to the processed T1-weighted images following the *dtiInit* pipeline from the VISTASOFT package (https://github.com/vistalab/vistasoft).

After preprocessing of the diffusion- and T1-weighted MRI data, we performed several quality checks on the processed data. First, we used FSL's *eddy QC* tool to perform quality assessment of diffusion data on the single-subject level (Bastiani et al., 2019). Based on the output of this tool, we excluded children with large head movement values, more than 2 SDs

from each independent sample's mean movement. Second, a diffusion tensor model was fitted at each voxel using FSL's *dtifit* and the eigenvectors of prominent tracts as the superior longitudinal fascicle, corpus callosum, and corticospinal tract were visually checked on the FA images. Third, successful brain extraction was visually checked, both for the diffusion- and T1-weighted MRI data.

Automated tractography pipeline

To reconstruct the left dorsal and ventral language-related fiber pathways and the corticospinal tract as a control from the preprocessed dMRI data, we used the python-based version of the Automated Fiber Quantification (pyAFQ 0.12) software (Kruper et al., 2021; Yeatman et al., 2012). This pipeline utilizes several steps to generate well-defined and anatomically plausible fiber pathways. Further, it results in global measures such as the streamline count of each pathway, as well as so-called tract profiles of local assessments of microstructural diffusion properties along each pathway (Kruper et al., 2021; Yeatman et al., 2012).

We performed the following steps as provided by the pyAFQ pipeline. The distribution of fiber orientation at each voxel was estimated using the Constrained Spherical Deconvolution (CSD) model (Tournier et al., 2004). A whole-brain probabilistic tractography was then performed with maximum turning angle = 30°, minimum streamline length = 10 mm, maximum length = 1000 mm, and step-size = 0.5 mm.

The resulting tractograms were then segmented into the three language-related fiber pathways of interest and the additional control tract by using predefined endpoint and waypoint ROIs, based on prior anatomical knowledge. For the endpoint ROIs, regions from the AAL atlas (Tzourio-Mazoyer et al., 2002) were selected to define the cortical termination points of each fiber pathway. Waypoint ROIs were defined in MNI space and then registered to children's native space using a non-linear transformation to further guide the segmentation to result in

anatomically plausible pathways (Wakana et al., 2007). Within the pyAFQ pipeline, endpoint and waypoint ROIs for over 20 major fiber pathways are provided, including the arcuate fascicle, IFOF, and corticospinal tract. To disentangle the dorsal pathway to BA44 from the one to BA6, we modified the default segmentation pipeline for the arcuate fascicle (see *Supplementary Figure 1*). For this, we used the default endpoint ROI in the temporal lobe but defined BA44 and BA6 as endpoint ROIs in the frontal cortex, respectively. Further, we used the default waypoint ROIs of the arcuate fascicle for both dorsal fiber pathways. For the dorsal pathway to BA6, we additionally defined an exclusion ROI in BA44 to disentangle streamlines of both dorsal fiber pathways, and a second exclusion ROI to ensure that no fibers would branch into the corticospinal tract. To segment the ventral fiber pathway, corresponding to the IFOF and the corticospinal tract as control tract, we relied on the default endpoint and waypoint ROIs provided by pyAFQ.

The segmented fiber pathways were then further refined by comparing each candidate fiber to a probability map (Hua et al., 2008) and discarding fibers passing low probability regions. By default, pyAFQ provides probability maps for the arcuate fascicle, which was used for both dorsal fiber pathways, for the IFOF, and corticospinal tract. Outlier fibers that deviated substantially from the core of the respective fiber pathway were additionally removed to result in a compact fiber bundle. To identify the fiber core, each pathway was clipped at the endpoint ROIs, resampled into 100 equidistant nodes along the pathway and the mean location of each node was calculated. The resulting bundles of each child were visually inspected to ensure biologically plausible identification of each fiber pathway. Fiber bundles of the two dorsal fiber pathways were excluded if streamlines followed anatomically implausible turns, or if the frontal endpoints of the bundles did not terminate in distinct areas of BA44 or BA6 (for examples see *Supplementary Figure 2*). The visual inspection led to the

exclusion of N = 4 of the dorsal pathway to BA44, N = 2 of the dorsal pathway to BA6, and N = 1 of the corticospinal tract in Sample 1, and N = 1 of the IFOF Sample 2.

Diffusion properties (i.e., fractional anisotropy [FA], mean diffusivity [MD], and radial diffusivity [RD]) were quantified at each node along each fiber pathway using a weighted approach as described by Yeatman et al. (2012). This resulted in a so-called tract profile that was used for further analysis. In addition, total streamline count of each fiber pathway was calculated.

We additionally reconstructed the homologous dorsal and ventral language pathways and the corticospinal tract as a control in children's right hemisphere. These exploratory analyses are reported in the Supplementary Material.

Statistical analysis

To estimate the relation between children's morpho-syntactic ability and structural properties of the language-related fiber pathways and the control tract, we fit multiple linear regressions using the lm() function from the stats package included in R (R Core Team, 2023) using RStudio (RStudio Team, 2020). As preregistered, we used the morpho-syntax scores as the predictor and the brain structural measures (i.e., total streamline count, FA, MD, and RD) as dependent variables in each model. We included both age groups in the analyses, with age group as factor and tested for a main effect and an interaction between age group and children's morpho-syntax scores. These analyses were performed independently in the two samples. Significant interactions with age group were followed up in each age group separately. In addition, for the purpose of comparison, we also explored language scores with a nonsignificant age interaction in the separate age groups. In all models, we controlled for sex, non-verbal IQ, handedness, and estimated intracranial volume (eTIV) to examine the specificity of the effects. As Sample 1 of the 3- and 5-year-olds consisted of children with and without a family history of dyslexia, we included family history as an additional covariate in

the analyses of this dataset. In all models testing for a main effect, we performed one-sided tests as streamline count and FA increase with maturation and higher cognitive function, while MD and RD decrease (Lebel & Deoni, 2018). Since streamline count constitutes one value per fiber pathway, the alpha-level was set at $p = 0.05$ for the models including streamline count. In the node-based approach (i.e., for FA, MD, and RD), multiple linear regressions were performed pointwise along 100 segments (i.e., nodes) of each fiber pathway. We therefore applied cluster-wise correction for multiple comparisons at $p = 0.05$ using non-parametric permutation-based correction (Nichols & Holmes, 2002) with a further Bonferroni correction for the N = 3 microstructural indices. Statistics for the cluster-wide relationship between these indices and morpho-syntax scores were calculated on the mean FA, MD, or RD of the respective significant cluster. We winsorized all language data and brain maturational indices limiting extreme values to the 95th percentile prior to conducting analyses.

To relate brain maturational indices of the language-related fiber pathways to preschooler's grammar performance on the sentence-level, we performed similar analyses on children's syntactic comprehension and production scores as reported in the Supplementary Material.

Results

To investigate the relation between preschooler's abilities in morpho-syntactic processing and brain structural maturation of language-related fiber pathways, we reconstructed two dorsal fiber pathways, one targeting BA44 and one targeting BA6, as well as the ventral pathway (i.e., IFOF), and the corticospinal tract as a control tract (see Fig. 1). In a preregistered procedure, we then related their brain maturational indices (i.e., total streamline count and local measures of FA, MD, and RD along 100 nodes of each pathway) to children's morpho-syntactic abilities, which were assessed using a noun plural assignment task. We examined this relation in two independent samples of 3- and 5-year-old (Sample 1), and 3- and 4-year-old (Sample 2) children. All effects were independent of sex, non-verbal IQ, handedness, eTIV, and, in Sample 1, additionally independent of family history of dyslexia. For models with streamline count, the significance threshold was set at $p = 0.05$. For node-based models, multiple comparison correction was applied at $p = 0.05$ with a further Bonferroni-correction to correct for N = 3 brain maturational indices.

In both samples, younger children at the age of 3 years differed from older children aged 4 and 5 years. In Sample 1, there was no significant main effect of children's morpho-syntax scores on their brain maturational measures in the reconstructed pathways, but a significant interaction with age group. This age interaction was found in the dorsal pathway to BA6 (anterior part: RD, node size = 15, node range = 17-31, $\beta = -0.03$, SE = 0.01, $t(78) = -3.076$, $p = 0.003$) and in the ventral pathway (streamline count, $\beta = 413.4$, SE = 181.5, $t(80) = 2.278$, $p = 0.025$; central part: FA, node size = 14, node range = 39-52, $\beta = 0.032$, SE = 0.001, $t(80) = 4.400$, $p < 0.001$; RD, node size = 12, node range = 40-51, $\beta = -0.031$, SE = 0.01, $t(80) = -3.166$, $p = 0.002$), but not the dorsal pathway to BA44. No main effect or interaction with age was found in the corticospinal tract serving as control tract. To follow-up on these interactions, we analyzed the two age groups, i.e. the 3- and 5-year-olds, separately. This

revealed that the interactions were driven by the 5-year-old children who showed a significant relation between their morpho-syntax scores and brain maturational measures in the dorsal pathway to BA6 (anterior part: FA, node size = 14, node range = 21-34, β = 0.0230, SE = 0.008, $t(47) = 3.069$, $p = 0.002$; RD, node size = 16, node range = 18-33, β = -0.0197, SE = 0.007, $t(47) = -2.832$, $p = 0.003$; see Fig. 2a) and the ventral pathway (streamline count, β = 322.0, SE = 136.3, $t(48) = 2.363$, $p = 0.011$; central part: FA, node size = 17, node range = 42-58, β = 0.021, SE = 0.065, $t(48) = 3.241$, $p = 0.001$). No effect was found in the dorsal pathway to BA44 or the control tract in the 5-year-olds. The 3-year-olds did not show any significant relation in either of the three language-related fiber pathways or the control tract.

In Sample 2, again no main effect of children's morpho-syntax scores on their language-related pathways was found. Further, we replicated the interaction with age group in the dorsal pathway to BA6 (part terminating in the temporal lobe: MD, node size = 28, node range = 73-100, β = -0.031, SE = 0.009, $t(22) = -3.568$, $p = 0.002$). Additionally, however, we found an age interaction in the dorsal pathway to BA44 (posterior part: MD, node size = 20, node range = 71-90, β = -0.031, SE = 0.01, $t(22) = -3.202$, $p = 0.004$; anterior part: RD, node size = 20, node range =18-37, β = -0.045, SE = 0.011, $t(22) = -4.058$, $p < 0.001$). No interaction with age was found in the ventral pathway, nor the control tract. Follow-up analyses separated by age group revealed that these interactions were again driven by the older age group. That is, the 4-year-old children, but not the 3-year-olds, showed a significant relation in the dorsal pathway to BA6 (anterior part: MD, node size = 20, node range = 10-29, β = -0.020, SE = 0.007, $t(11) = -2.965$, $p = 0.006$; see Fig. 2b) and in the dorsal pathway to BA44 (anterior part: MD, node size = 27, node range = 7-33, β = -0.016, SE = 0.004, $t(11) = -4.172$, $p < 0.001$; RD, node size = 11, node range = 27-37, β = -0.023, SE = 0.007, $t(11) = -3.240$, $p = 0.004$). In line with the absence of an interaction, no significant effect was found in the ventral pathway or the corticospinal tract in either of the two age groups.

In order to test whether these correlations between behavior and white matter fiber tracts were specific for the left hemisphere, we also analyzed the corresponding white matter fiber tracts in the right hemisphere. For the right-hemispheric homologous fiber pathways, we found no main effect, but a significant age interaction driven by the older age groups in the right ventral pathway in Sample 1 and in the right dorsal pathway to BA6 in Sample 2 for children's morpho-syntax scores (see *Supplementary Results* for more details). No interaction or main effect was found for the right dorsal pathway to BA44 in either sample.

Further, to test whether the interaction with age found in our analyses could be due to the fact that the noun plural assignment task in 3-year-olds only contained real nouns, while in 4- and 5-year-olds additionally contained pseudo nouns, an analysis only including the real nouns was conducted. This was done as the generation of plural forms of real nouns and pseudo nouns might be based on different strategies since the plural forms of pseudo nouns cannot be retrieved from the mental lexicon but requires rule application. Thus, in an exploratory analysis, we divided the items presented to the 4- and 5-year-olds into real and pseudo nouns. The summed score of real noun items were standardized within both age groups in each sample independently, i.e. the 3- and 5-year-olds of Sample 1, and the 3- and 4-year-olds of Sample 2. We then tested for a relation between children's real noun morpho-syntax scores and maturational indices in the language pathways and the control tract.

For children's real noun morpho-syntax scores, as before, no significant main effect but significant interactions with age group were found in both samples. In Sample 1, the interaction was found in the dorsal pathway to BA6 (anterior part: MD, node size = 20, node range = 7-26, β = -0.025, SE = 0.007, $t(77) = -3.498$, $p < 0.001$; RD, node size = 18, node range =16-33, β = -0.041, SE = 0.011, $t(77) = -3.781$, $p < 0.001$), in the dorsal pathway to BA44 (anterior to central part: FA, node size = 9, node range = 29-37, β = 0.033, SE = 0.01, $t(75) = 3.171$, $p = 0.002$; RD, node size = 17, node range =28-44, β = -0.041, SE = 0.012,

$t(75) = -3.456$, $p < 0.001$), and the ventral pathway (streamline count, $\beta = 453.4$, SE = 213.0, $t(79) = 2.128$, $p = 0.036$; central part: FA, node size = 13, node range =40-52, $\beta = 0.034$, SE = 0.009, $t(79) = 3.762$, $p < 0.001$). As before, follow-up analyses revealed that the present interactions were driven by the older age group. The 5-year-olds, but not the 3-year-olds, showed an effect in the dorsal pathway to BA6 (anterior part: RD, node size = 15, node range = 18-32, $\beta = -0.03$, SE = 0.009, $t(47) = -3.348$, $p < 0.001$) and in the dorsal pathway to BA44 (anterior to central part: RD, node size = 20, node range = 25-44, $\beta = -0.028$, SE = 0.009, $t(45) = -3.045$, $p = 0.002$; see Fig. 3a). No effect was found for the ventral pathway in either of the two age groups.

In Sample 2, the interaction was also located in the dorsal pathway to BA6 (anterior to central part: FA, node size = 11, node range = 35-45, $\beta = 0.075$, SE = 0.018, $t(22) = 4.121$, $p < 0.001$; MD, node size = 34, node range = 11-44, $\beta = -0.045$, SE = 0.013, $t(22) = -3.368$, $p = 0.003$; RD, node size = 14, node range =32-45, $\beta = -0.073$, SE = 0.016, $t(22) = -4.518$, $p < 0.001$) replicating the results of Sample 1. No significant interaction was found in the dorsal pathway to BA44 or the ventral pathway in this sample. Following-up on this interaction, 4-year-old children did not show a significant effect in the dorsal pathway to BA6, but a significant relation in the dorsal pathway to BA44 (anterior part: MD, node size = 17, node range = 27-43, $\beta = -0.026$, SE = 0.008, $t(11) = -3.296$, $p = 0.004$; see Fig. 3b). No effect was found for the ventral pathway. No significant effect for the real noun morpho-syntax scores was found in the 3-year-olds in either sample. Moreover, no main effect or interaction was found for the control tract. These results indicate that the age difference between the 3-year-olds and the older children found for the entire set of nouns was not driven by different subsets of the experimental material.

Analyses for the pseudo noun plurals can be found in the Supplementary material.

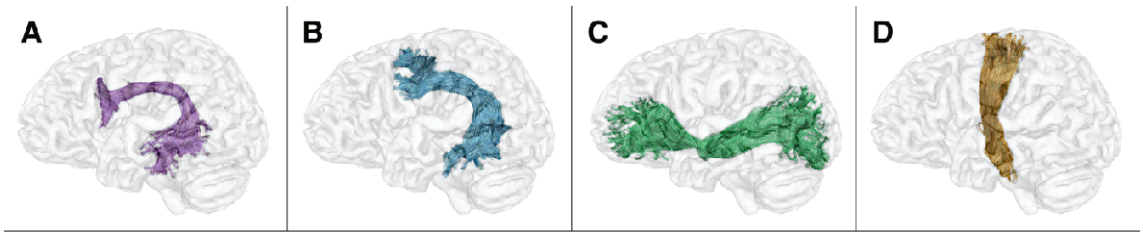

Figure 1: Example of the segmented and refined language fiber pathways – a) the dorsal pathway to BA44 (purple), b) the dorsal pathway to BA6 (blue), the c) ventral pathway (green) – and d) the corticospinal tract as control tract (orange).

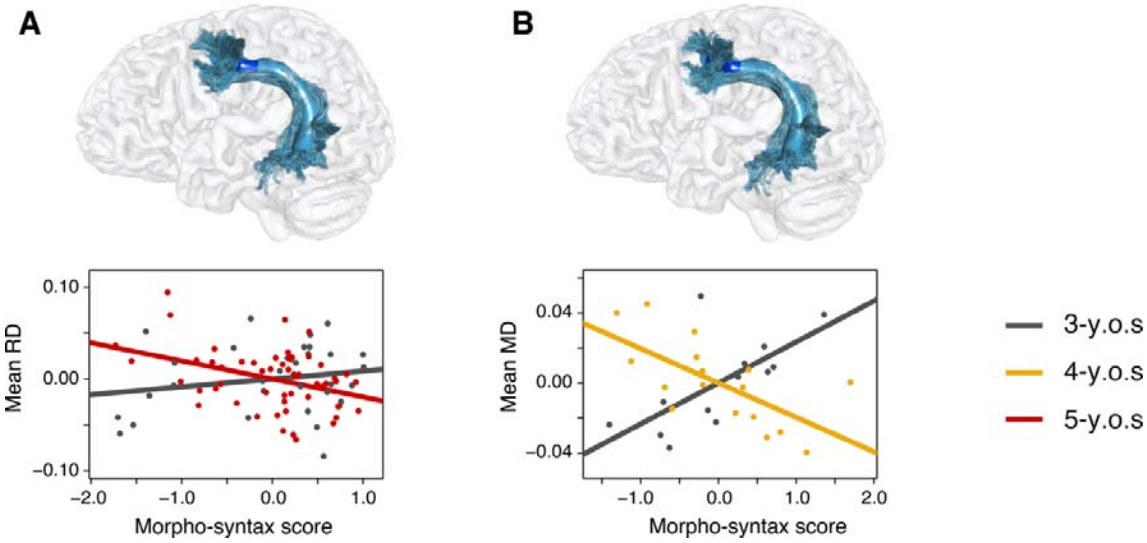

Figure 2: BA6 targeting dorsal pathway. Linear partial correlation of preschooler's morpho-syntactic abilities and brain structural measures within the significant clusters of the dorsal pathway to BA6 (blue) in (A) Sample 1 and (B) Sample 2. Cluster-wise correction was applied at p = 0.05 with a further Bonferroni-correction of N = 3. Residuals of the partial correlation of the morpho-syntax score and the mean values of the brain maturational measures within the significant clusters are plotted. (A) Significant relation of 5-year-old's radial diffusivity (RD) and morpho-syntax scores in Sample 1 (3-y.o.s: r_p = 0.19, p = 0.33, gray; 5-y.o.s: r_p = -0.38, p = 0.007, red). (B) Significant relation of 4-year-old's mean diffusivity (MD) and morpho-syntax scores in Sample 2 (3-y.o.s: r_p = 0.67, p = 0.05, gray; 4-y.o.s: r_p = -0.67, p = 0.013, yellow).

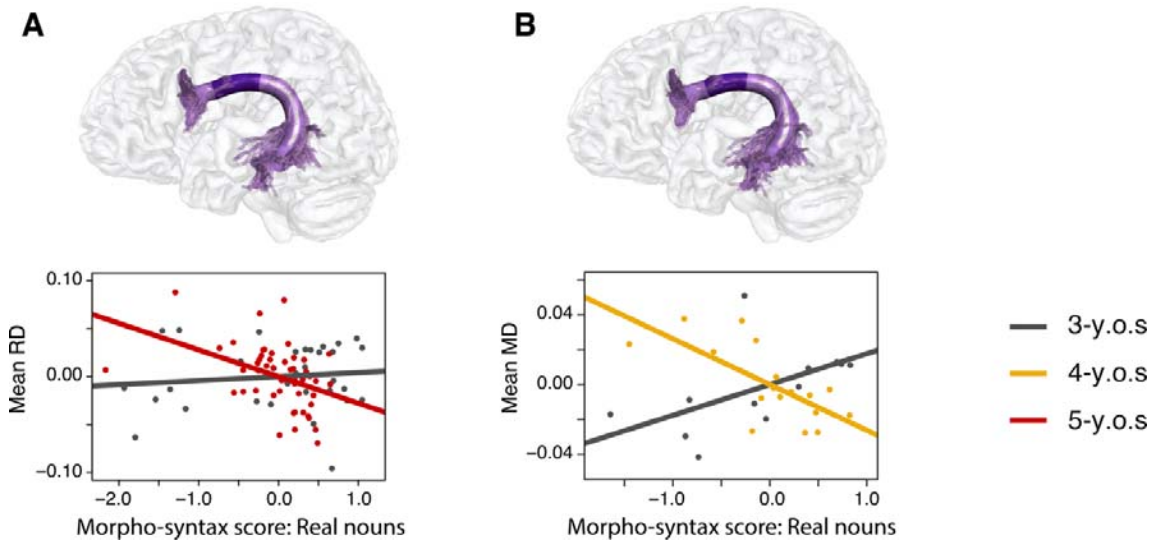

Figure 3: BA44 targeting dorsal pathway. Linear partial correlation of preschooler's morpho-syntactic abilities tested with real nouns and brain structural measures within the significant clusters of the dorsal pathway to BA44 (purple) in (A) Sample 1 and (B) Sample 2. Cluster-wise correction was applied at p = 0.05 with a further Bonferroni-correction of N = 3. Residuals of the partial correlation of the morpho-syntax score tested with real nouns and the mean values of the brain maturational measures within the significant clusters are plotted. (A) Significant relation between 5-year-old's real noun morpho-syntax scores and radial diffusivity (RD) in Sample 1 (3-y.o.s: r_p = 0.12, p = 0.59, gray; 5-y.o.s: r_p = -0.41, p = 0.004, red).(B) Significant relation between 4-year-old's real noun morpho-syntax scores and mean diffusivity (MD) in Sample 2 (3-y.o.s: r_p = 0.61, p = 0.08, gray; 4-y.o.s: r_p = -0.70, p = 0.007, yellow).

Discussion

The present paper investigated the relation between grammar acquisition at the morpho-syntactic level and brain maturation at preschool age. Children between 3 and 5 years of age are known to undergo major behavioral improvements in their morpho-syntactic performance (Kauschke, 2012). However, the dorsal fiber pathway to BA44 supporting grammar in adults matures relatively late during development (Brauer et al., 2013; Skeide & Friederici, 2016). This raises the question of whether this fiber pathway already supports morpho-syntax acquisition in the early preschool years, or whether earlier maturing brain structures take on this role. We investigated this question in two independent samples of 3- and 5-year-old children (Sample 1), and 3- and 4-year-old children (Sample 2). Our analysis showed consistent differences in the association between the maturation of language fiber pathways and children's morpho-syntactic abilities between the 3-year-olds and the older age groups. In both samples, the older age group showed a relation of their morpho-syntax scores with the anterior part of the dorsal pathway to BA6. In an exploratory analysis, in which we analyzed this relation in real nouns only, we further found a consistent additional relation of the dorsal pathway to BA44 with morpho-syntax abilities in the 4- and 5-year-olds, but not in the 3-year-olds. These results reveal a clear difference of the involvement of the dorsal pathways BA6 and BA44 in the older compared to the younger children.

In the adult brain, the dorsal pathway to BA6 has been found to support auditory-to-motor mapping and phonological processes as in speech repetition (Saur et al., 2008). In the developing brain, this fiber pathway is already detectable in newborns, suggesting a relatively mature white matter structure at birth (Perani et al., 2011). Thus, it was hypothesized that early language acquisition might be primarily supported by the dorsal pathway to BA6 (Skeide & Friederici, 2016). Our findings that children's morpho-syntactic abilities, which reflect rule-based processes at the word-level, were associated with this fiber pathway are in

line with this hypothesis. Further, the dorsal pathway to BA6 might support the detection of statistical regularities in phonological sequences. From behavioral studies, we know that statistical learning is crucial for language acquisition in general and for grammar development in particular (Saffran, 2020). The German plural rule system relies on the phonology of a given noun, although not exclusively (Köpcke, 1988). The present association with the dorsal pathway to BA6 with children's morpho-syntactic abilities might point to a phonology-based statistical acquisition strategy. Thus, in contrast to adults, children might strongly rely on learning statistical regularities to acquire grammatical rules, be it on the sentence-level (Friederici et al., 2011) or the word-level as in the present study, supported by the dorsal pathway to BA6.

In addition, we found a consistent relation between 4- and 5-year-olds' real noun morpho-syntax scores and the dorsal pathway to BA44, which was also observed for the 4-year-old's overall morpho-syntax scores. In the mature brain, this pathway supports rule-based language processes, as found both in adult and patient studies (Friederici & Gierhan, 2013; Griffiths et al., 2013; Wilson et al., 2011). In children, the maturation of the dorsal pathway to BA44 is related to sentence processing abilities from 4 years of age (Skeide et al., 2016). Preschool children evidently build up a rule system during the acquisition of plural nouns (Kauschke, 2012). This is apparent in the overgeneralization errors children produced that mainly follow frequent plural rules (see *Supplementary Figure 3*; Kauschke et al., 2011; Laaha, 2011; Szagun, 2001). Thus, maturation of the dorsal pathway to BA44 might not only support the acquisition of complex sentences (Skeide et al., 2016) but also of rule-based processes on the word-level.

To summarize, we found a consistent association between 4- and 5-year-old children's abilities to apply morpho-syntactic rules and the maturation of both dorsal pathways (to BA44 and to BA6) across two independent samples. In adults, the dorsal connection between BA44

and the temporal cortex is crucial for rule-based processes, such as morpho-syntax (Meykadeh et al., 2023; Rolheiser et al., 2011). Against this background, our findings suggest that children may employ both statistical and rule-based learning strategies to acquire grammar in the early preschool years, supported by the maturation of the dorsal fiber pathways to BA6 and BA44.

Limitations

While the main findings were consistent across two independent samples – the association of the dorsal pathway to BA6 with children's morpho-syntax scores and the dorsal pathway to BA44 for real noun morpho-syntax scores – a few additional findings were observed in only one of the two investigated samples. This highlights the need for replication of these findings in larger samples, which will be a significant challenge for future research given the difficulty of acquiring MRI data in young preschool children under 5 years of age. This would also help clarify which structures are involved in grammar before the age of 4. Further, in the present study we investigated grammar ability at the word-level, i.e. morpho-syntax. An exciting avenue for future research is the question of which age the association between grammar ability on the sentence-level and brain structural maturation of the dorsal pathway to BA44 arises, given the ongoing maturation of this pathway even during middle childhood (Brauer et al., 2013; Skeide et al., 2016). Moreover, here we tested children's morpho-syntactic abilities with nouns. Given the different trajectories of nouns and verbs in language acquisition and the findings of dissociative neural correlates in adults for nouns versus verbs (Vigliocco et al., 2011; Waxman et al., 2013), future research should investigate whether the present findings also hold for the morpho-syntactic inflection of verbs.

Conclusion

In sum, the present results highlight the role of the white matter language network for morpho-syntax acquisition during the critical preschool years between 3 and 5 years of age.

The 4- and 5-year-olds, but not the 3-year-olds, showed a relationship between morpho-syntax and both the dorsal pathway to BA44 and the dorsal pathway to BA6. While in adults, the dorsal pathway to BA44 supports rule-based and the dorsal pathway to BA6 supports phonological processes, preschool-aged children seem to rely on both dorsal fiber pathways for morpho-syntax. Our findings suggest that young children may employ different strategies for acquiring the rule-based grammar system. The dorsal pathway to BA6 may support statistical learning mechanisms, whereas the maturation of the dorsal pathway to BA44 may support the acquisition of a robust rule-based grammar system.

References

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20*(2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

Bastiani, M., Cottaar, M., Fitzgibbon, S. P., Suri, S., Alfaro-Almagro, F., Sotiropoulos, S. N., Jbabdi, S., & Andersson, J. L. R. (2019). Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction. *NeuroImage*, *184*, 801–812. https://doi.org/10.1016/j.neuroimage.2018.09.073

Becker, Y., Loh, K. K., Coulon, O., & Meguerditchian, A. (2022). The Arcuate Fasciculus and language origins: Disentangling existing conceptions that influence evolutionary accounts. *Neuroscience & Biobehavioral Reviews*, *134*, 104490. https://doi.org/10.1016/j.neubiorev.2021.12.013

Berger, P., Friederici, A. D., & Wiesmann, C. G. (2021). *Maturation of distinct neural components of the cognitive control network support early development of inhibitory control*. https://doi.org/10.1101/2021.07.02.450852

Berko, J. (1958). The Child's Learning of English Morphology. WORD, *14*(2–3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Brauer, J., Anwander, A., Perani, D., & Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain and Language*, *127*(2), 289–295. https://doi.org/10.1016/j.bandl.2013.03.001

Bulut, T. (2022). Neural correlates of morphological processing: An activation likelihood estimation meta-analysis. *Cortex*, *151*, 49–69. https://doi.org/10.1016/j.cortex.2022.02.010

Crain, S., & Thornton, R. (2012). Syntax acquisition. *WIREs Cognitive Science*, *3*(2), 185–203. https://doi.org/10.1002/wcs.1158

Eichner, C., Berger, P., Klein, C. C., & Friederici, A. D. (2024). Lateralization of dorsal fiber tract targeting Broca's area concurs with language skills during development. *Progress in Neurobiology*, *236*, 102602. https://doi.org/10.1016/j.pneurobio.2024.102602

Fan, Q., Eichner, C., Afzali, M., Mueller, L., Tax, C. M. W., Davids, M., Mahmutovic, M., Keil, B., Bilgic, B., Setsompop, K., Lee, H.-H., Tian, Q., Maffei, C., Ramos-Llordén, G., Nummenmaa, A., Witzel, T., Yendiki, A., Song, Y.-Q., Huang, C.-C., … Huang, S. Y. (2022). Mapping the human connectome using diffusion MRI at 300 mT/m gradient strength: Methodological advances and scientific impact. *NeuroImage*, *254*, 118958. https://doi.org/10.1016/j.neuroimage.2022.118958

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., Miller, K. L., Ugurbil, K., & Yacoub, E. (2010). Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI and Fast Diffusion Imaging. *PLoS ONE*, *5*(12), e15710. https://doi.org/10.1371/journal.pone.0015710

Fernández-Miranda, J. C., Wang, Y., Pathak, S., Stefaneau, L., Verstynen, T., & Yeh, F.-C. (2015). Asymmetry, connectivity, and segmentation of the arcuate fascicle in the human brain. *Brain Structure and Function*, *220*(3), 1665–1680. https://doi.org/10.1007/s00429-014-0751-7

Fox, D., & Grodzinsky, Y. (1998). Children's Passive: A View from the *By* -Phrase.

Linguistic Inquiry, *29*(2), 311–332. https://doi.org/10.1162/002438998553761

Fridriksson, J., Yourganov, G., Bonilha, L., Basilakos, A., Den Ouden, D.-B., & Rorden, C. (2016). Revealing the dual streams of speech processing. *Proceedings of the National Academy of Sciences*, *113*(52), 15108–15113. https://doi.org/10.1073/pnas.1614038114

Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*(2), 250–254. https://doi.org/10.1016/j.conb.2012.10.002

Friederici, A. D., Mueller, J. L., & Oberecker, R. (2011). Precursors to Natural Grammar Learning: Preliminary Evidence from 4-Month-Old Infants. *PLoS ONE*, *6*(3), e17920. https://doi.org/10.1371/journal.pone.0017920

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *NeuroImage*, *114*, 294–302. https://doi.org/10.1016/j.neuroimage.2015.04.011

Griffiths, J. D., Marslen-Wilson, W. D., Stamatakis, E. A., & Tyler, L. K. (2013). Functional Organization of the Neural Language System: Dorsal and Ventral Pathways Are Critical for Syntax. *Cerebral Cortex*, *23*(1), 139–147. https://doi.org/10.1093/cercor/bhr386

Grimm, H. (2001). *Sprachentwicklungstest für drei- bis fünfjährige Kinder: SETK 3–5. Diagnose von Sprachverarbeitungsfähigkeiten und auditiven Gedächtnisleistungen*. Hogrefe: Verlag für Psychologie.

Grodzinsky, Y., Pieperhoff, P., & Thompson, C. (2021). Stable brain loci for the processing of complex syntax: A review of the current neuroimaging evidence. *Cortex*, *142*, 252–271. https://doi.org/10.1016/j.cortex.2021.06.003

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Developmental Science*, *20*(5), e12445. https://doi.org/10.1111/desc.12445

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2020). Two systems for thinking about others' thoughts in the developing brain. *Proceedings of the National Academy of Sciences*, *117*(12), 6928–6935. https://doi.org/10.1073/pnas.1916725117

Grosse Wiesmann, C., Schreiber, J., Singer, T., Steinbeis, N., & Friederici, A. D. (2017). White matter maturation is associated with the emergence of Theory of Mind in early childhood. *Nature Communications*, *8*(1), 14692. https://doi.org/10.1038/ncomms14692

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. MIT Press.

Harley, H. (2015). 33. The Syntax-Morphology Interface. In T. Kiss & A. Alexiadou (Eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 42/2* (pp. 1128–1154). DE GRUYTER. https://doi.org/10.1515/9783110363708-010

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1–2), 67–99. https://doi.org/10.1016/j.cognition.2003.10.011

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., Calabresi, P. A., Pekar, J. J., Van Zijl, P. C. M., & Mori, S. (2008). Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, *39*(1), 336–347. https://doi.org/10.1016/j.neuroimage.2007.07.053

Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. DE GRUYTER. https://doi.org/10.1515/9783110283891

Kauschke, C., Kurth, A., & Domahs, U. (2011). Acquisition of German Noun Plurals in

Typically Developing Children and Children with Specific Language Impairment. *Child Development Research*, *2011*, 1–17. https://doi.org/10.1155/2011/718925

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs□ringing artifact removal based on local subvoxel□shifts. *Magnetic Resonance in Medicine*, *76*(5), 1574–1581. https://doi.org/10.1002/mrm.26054

Klein, C. C., Berger, P., Goucha, T., Friederici, A. D., & Grosse Wiesmann, C. (2023). Children's syntax is supported by the maturation of BA44 at 4 years, but of the posterior STS at 3 years of age. *Cerebral Cortex*, *33*(9), 5426–5435. https://doi.org/10.1093/cercor/bhac430

Köpcke, K.-M. (1988). Schemas in German plural formation. *Lingua*, *74*(4), 303–335. https://doi.org/10.1016/0024-3841(88)90064-2

Kruper, J., Yeatman, J. D., Richie-Halford, A., Bloom, D., Grotheer, M., Caffarra, S., Kiar, G., Karipidis, I. I., Roy, E., Chandio, B. Q., Garyfallidis, E., & Rokem, A. (2021). Evaluating the Reliability of Human Brain White Matter Tractometry. *Aperture Neuro*, *2021*(1), 25. https://doi.org/10.52294/e6198273-b8e3-4b63-babb-6e6b0da10669

Kuhl, U., Neef, N. E., Kraft, I., Schaadt, G., Dörr, L., Brauer, J., Czepezauer, I., Müller, B., Wilcke, A., Kirsten, H., Emmrich, F., Boltze, J., Friederici, A. D., & Skeide, M. A. (2020). The emergence of dyslexia in the developing brain. *NeuroImage*, *211*, 116633. https://doi.org/10.1016/j.neuroimage.2020.116633

Kuhl, U., Sobotta, S., Legascreen Consortium, & Skeide, M. A. (2021). Mathematical learning deficits originate in early childhood from atypical development of a frontoparietal brain network. *PLOS Biology*, *19*(9), e3001407. https://doi.org/10.1371/journal.pbio.3001407

Laaha, S. (2011). Sonority, gender and the impact of suffix predictability on the acquisition of German noun plurals. *Language, Interaction and Acquisition*, *2*(1), 82–100. https://doi.org/10.1075/lia.2.1.04laa

Lebel, C., & Deoni, S. (2018). The development of brain white matter microstructure. *NeuroImage*, *182*, 207–218. https://doi.org/10.1016/j.neuroimage.2017.12.097

Leminen, A., Smolka, E., Duñabeitia, J. A., & Pliatsikas, C. (2019). Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex*, *116*, 4–44. https://doi.org/10.1016/j.cortex.2018.08.016

Liebig, J., Friederici, A. D., Neef, N. E., Friederici, A. D., Emmrich, F., Brauer, J., Wilcke, A., Neef, N. E., Boltze, J., Skeide, M., Kirsten, H., Schaadt, G., Müller, B., Kraft, I., Czepezauer, I., & Dörr, L. (2020). Auditory brainstem measures and genotyping boost the prediction of literacy: A longitudinal study on early markers of dyslexia. *Developmental Cognitive Neuroscience*, *46*, 100869. https://doi.org/10.1016/j.dcn.2020.100869

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, *283*(5398), 77–80. https://doi.org/10.1126/science.283.5398.77

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.-F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage*, *49*(2), 1271–1281. https://doi.org/10.1016/j.neuroimage.2009.10.002

Melchers, P., & Preuß, U. (2003). *K-ABC Kaufman Assessment Battery for Children, Deutsche Version Individualtest zur Messung von Intelligenz und Fertigkeit bei Kindern (6. Teilweise ergänzte Aufl.)*. Hogrefe.

Meykadeh, S., Khadem, A., Sulpizio, S., & Sommer, W. (2023). Functional connectivity during morphosyntactic processing: An fMRI study in balanced Turkish-Persian

bilinguals. *Journal of Neurolinguistics*, *68*, 101162. https://doi.org/10.1016/j.jneuroling.2023.101162

Morelli, J. N., Runge, V. M., Feiweier, T., Kirsch, J. E., Williams, K. W., & Attenberger, U. I. (2010). Evaluation of a Modified Stejskal-Tanner Diffusion Encoding Scheme, Permitting a Marked Reduction in TE, in Diffusion-Weighted Imaging of Stroke Patients at 3 T: *Investigative Radiology*, *45*(1), 29–35. https://doi.org/10.1097/RLI.0b013e3181c65c11

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. https://doi.org/10.1002/hbm.1058

Oberecker, R., & Friederici, A. D. (2006). Syntactic event-related potential components in 24-month-olds' sentence comprehension. *NeuroReport*, *17*(10), 1017–1021. https://doi.org/10.1097/01.wnr.0000223397.12694.9a

Oberecker, R., Friedrich, M., & Friederici, A. D. (2005). Neural Correlates of Syntactic Processing in Two-Year-Olds. *Journal of Cognitive Neuroscience*, *17*(10), 1667–1678. https://doi.org/10.1162/089892905774597236

O'Brien, K. R., Kober, T., Hagmann, P., Maeder, P., Marques, J., Lazeyras, F., Krueger, G., & Roche, A. (2014). Robust T1-Weighted Structural Brain Imaging and Morphometry at 7T Using MP2RAGE. *PLoS ONE*, *9*(6), e99676. https://doi.org/10.1371/journal.pone.0099676

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Peña, E. A., & Slate, E. H. (2006). Global Validation of Linear Model Assumptions. *Journal of the American Statistical Association*, *101*(473), 341–354. https://doi.org/10.1198/016214505000000637

Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., & Friederici, A. D. (2011). Neural language networks at birth. *Proceedings of the National Academy of Sciences*, *108*(38), 16056–16061. https://doi.org/10.1073/pnas.1102991108

Qi, T., Schaadt, G., & Friederici, A. D. (2021). Associated functional network development and language abilities in children. *NeuroImage*, *242*, 118452. https://doi.org/10.1016/j.neuroimage.2021.118452

R Core Team. (2023). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing* [Computer software]. https://www.R-project.org/

Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic Processing in the Human Language System: Synergy between the Arcuate Fascicle and Extreme Capsule. *Journal of Neuroscience*, *31*(47), 16949–16957. https://doi.org/10.1523/JNEUROSCI.2725-11.2011

RStudio Team. (2020). *Rstudio: Integrated development environment for r* [Computer software]. RStudio, PBC. http://www.rstudio.com/

Saffran, J. R. (2020). Statistical Language Learning in Infancy. *Child Development Perspectives*, *14*(1), 49–54. https://doi.org/10.1111/cdep.12355

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035–18040. https://doi.org/10.1073/pnas.0805234105

Schipke, C. S., Knoll, L. J., Friederici, A. D., & Oberecker, R. (2012). Preschool children's interpretation of object□initial sentences: Neural correlates of their behavioral

performance. *Developmental Science*, *15*(6), 762–774. https://doi.org/10.1111/j.1467-7687.2012.01167.x

Skeide, M. A., Brauer, J., & Friederici, A. D. (2014). Syntax gradually segregates from semantics in the developing brain. *NeuroImage*, *100*, 106–111. https://doi.org/10.1016/j.neuroimage.2014.05.080

Skeide, M. A., Brauer, J., & Friederici, A. D. (2016). Brain Functional and Structural Predictors of Language Performance. *Cerebral Cortex*, *26*(5), 2127–2139. https://doi.org/10.1093/cercor/bhv042

Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nature Reviews Neuroscience*, *17*(5), 323–332. https://doi.org/10.1038/nrn.2016.23

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Strotseva-Feinschmidt, A., Schipke, C. S., Gunter, T. C., Brauer, J., & Friederici, A. D. (2019). Young children's sentence comprehension: Neural correlates of syntax-semantic competition. *Brain and Cognition*, *134*, 110–121. https://doi.org/10.1016/j.bandc.2018.09.003

Szagun, G. (2001). Learning different regularities: The acquisition of noun plurals by German-speaking children. *First Language*, *21*(62), 109–141. https://doi.org/10.1177/014272370102106201

Tomasello, M., & Brooks, P. (1999). Early syntactic development: A Construction Grammar approach. In M. Barrett (Ed.), *The development of language* (pp. 161–190). Psychology Press.

Tournier, J.-D., Calamante, F., Gadian, D. G., & Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, *23*(3), 1176–1185. https://doi.org/10.1016/j.neuroimage.2004.07.037

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*, 116137. https://doi.org/10.1016/j.neuroimage.2019.116137

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, *2*(10), 717–726. https://doi.org/10.1038/35094573

Veraart, J., Fieremans, E., & Novikov, D. S. (2016). Diffusion MRI noise mapping using random matrix theory. *Magnetic Resonance in Medicine*, *76*(5), 1582–1593. https://doi.org/10.1002/mrm.26059

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, *35*(3), 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007

Vos, S. B., Tax, C. M. W., Luijten, P. R., Ourselin, S., Leemans, A., & Froeling, M. (2017). The importance of correcting for signal drift in diffusion MRI. *Magnetic Resonance*

in Medicine, *77*(1), 285–299. https://doi.org/10.1002/mrm.26124

Wakana, S., Caprihan, A., Panzenboeck, M. M., Fallon, J. H., Perry, M., Gollub, R. L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., Van Zijl, P., & Mori, S. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*, *36*(3), 630–644. https://doi.org/10.1016/j.neuroimage.2007.02.049

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. (2013). Are Nouns Learned Before Verbs? Infants Provide Insight Into a Long Standing Debate. *Child Development Perspectives*, *7*(3), 155–159. https://doi.org/10.1111/cdep.12032

Wechsler, D. (2009). *Wechsler Preschool and Primary Scale of Intelligence-III, deutsche Version (WPPSI-III).* [Computer software].

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., Ogar, J. M., DeLeon, J., Miller, B. L., & Gorno-Tempini, M. L. (2011). Syntactic Processing Depends on Dorsal Language Tracts. *Neuron*, *72*(2), 397–403. https://doi.org/10.1016/j.neuron.2011.09.014

Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., & Feldman, H. M. (2012). Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. *PLoS ONE*, *7*(11), e49790. https://doi.org/10.1371/journal.pone.0049790

Zaccarella, E., Meyer, L., Makuuchi, M., & Friederici, A. D. (2017). Building by Syntax: The Neural Basis of Minimal Linguistic Structures. *Cerebral Cortex*, *27*(1), 411–421. https://doi.org/10.1093/cercor/bhv234

Acknowledgements

We thank Gesa Schaadt, Michael Skeide, Ting Qi and the entire LEGASCREEN consortium, and Hung Nguyen Trong and Christiane Attig for their support in data acquisition and sourcing. We thank Cornelius Eichner and Alfred Anwander for their methodological advice. We further thank Kerstin Flake, Andrea Gast-Sandmann and Heike Schmidt-Duderstedt for help with the figures, and Joshua Grant for helpful comments to a previous version of this manuscript. Further, we thank the Max Planck Förderstiftung for granting a writing stipend to C.C.K.

Author Contributions

Cheslie C. Klein: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Formal analysis, Writing – Original Draft. **Philipp Berger**: Data curation, Methodology, Software, Supervision, Writing – Review & Editing. **Charlotte Grosse Wiesmann**: Conceptualization, Funding acquisition, Resources, Supervision, Investigation, Writing – Review & Editing. **Angela D. Friederici**: Conceptualization, Funding acquisition, Resources, Supervision, Writing – Review & Editing.

Funding

This study was funded by the European Research Council (ERC) Starting Grant (project number REPRESENT 101117806) to CGW.

Competing Interests

The authors declare no competing interests.

Data availability statement

All employed materials and data discussed in the paper are saved in a local repository at the Max Planck Institute for Human Cognitive and Brain Sciences, and data in fully anonymized format will be made available to the reader upon request (according to the data protection policy in the ethics agreement).