The Relationship between Memory-Related White Matter Tracts and Second Language Grammar Learning

Suzanne C. op ‘t Hof, BSc.

Donders Institute

Abstract

An important aspect of learning another language is grammar learning, which depends on the declarative memory system. Three white matter tracts are associated with declarative memory, namely the uncinate fasciculus, arcuate fasciculus, and the white matter connection between the hippocampus and the medial prefrontal cortex. This study investigated the relationship between these three bilateral declarative memory-related white matter tracts and second language (L2) grammar learning. Participants learned Icelandic grammar for five consecutive days and were tested before and after learning on a grammatical judgement task in the magnetic resonance imaging (MRI) scanner. Diffusion weighted imaging (DWI) data was collected and the three memory-related white matter tracts were delineated using XTRACT. In total, 29 Dutch non-bilingual students were included for data-analyses. The results showed no relationship between the fractional anisotropy (FA) and mean diffusivity (MD) of the three memory-related white matter tracts and L2 grammar learning. More research is needed to fully support or reject this finding. Altogether, these findings suggest that memory-related white matter tracts are not associated with L2 grammar learning.

Keywords: *second language (L2) grammar learning; declarative memory; uncinate fasciculus; arcuate fasciculus; white matter connection between the hippocampus and medial prefrontal cortex; fractional anisotropy (FA); mean diffusivity (MD); diffusion tensor imaging (DTI)*

The Relationship between Memory-Related White Matter Tracts and Second Language Grammar Learning

Grammar learning is an important aspect of learning a second language. Many people, including adults, will learn a second language in their lives. In 2016, more than 80% of people with a university or college degree in the European Union knew at least one foreign language (Eurostat, 2019). Grammar learning is affected by several factors, namely level of education, motivation, learning strategies, white matter integrity, and (previous) experience with the language (Flöel, de Vries, Scholz, Breitenstein, & Johansen-Berg, 2009; He, 2013; Lahmann, Steinkrauss, & Schmid, 2016; Unsworth, Persson, Prins, & De Bot, 2015). The more experience someone has with a second language, the better the second language grammatical knowledge of that person is (Flege & Liu, 2001). Prior second language learning experience also makes learning another language easier (Golonka, 2010). This is true for both adults who acquire the second language during childhood, adolescence, or adulthood. Therefore, second language proficiency is an important factor of learning another language and is thus important for grammar learning. In addition to language proficiency, memory is essential for learning a second language (Zidane, 2016). Learning another language depends on declarative and procedural memory, the essential learning and memory systems of the brain (Ullman & Lovelett, 2018).

According to the declarative/procedural model, procedural memory is more involved in first language grammar learning, while declarative memory is more involved in second language (L2) grammar learning (Ullman, 2001). The declarative/procedural model posits that in a first language, the learning and use of grammar depends largely on procedural memory, whereas the learning and use of words depends largely on declarative memory. In contrast, when learning a second language as an adult, the learning and use of grammar is shifted towards declarative memory instead of procedural memory (Ullman, 2001). Procedural memory is the learning and the control of motor and cognitive skills and habits (Ullman, 2001). Declarative memory consists of episodic memory (remembering specific events of the past) together with semantic memory (conceptual knowledge) (Roediger, Zaromb, & Lin, 2017). The processes of declarative memory rely on the hippocampus and other related structures in the medial temporal lobe (Roediger et al., 2017). Hamrick (2015) investigated how declarative and procedural memory can predict L2 grammar learning and retention. Participants performed declarative and procedural memory tests and a syntactic semi-artificial language learning task, followed by an immediate and delayed recognition task about the sentences of this semi-artificial language learning task. The results show that second language syntax learning was correlated positive with declarative memory, and negatively with procedural memory in the immediate test. So, declarative memory as opposed to procedural memory is important for grammar learning (Hamrick, 2015). Declarative memory and grammar learning are both related to white matter structures in the brain (Flöel et al., 2009; Mabbott, Rovet, Noseworthy, Smith, & Rockel, 2009).

Learning performance is associated with white matter tissue properties (Huber, Donnelly, Rokem, & Yeatman, 2018). Variation in white matter structure, e.g. differences in tissue properties, correlates with individual differences in skill and ability performances (Johansen-Berg, 2010). These relationships have been found for a series of cognitive domains, including the language and memory domain. Johansen-Berg (2010) also explains that correlations between white matter structure and behavioral differences are local effects, rather than being related to global white matter structure in the brain. Therefore, we are interested in studying the relationship between individual variation in L2 grammar learning and white matter pathways.

Three white matter tracts that are related to declarative memory, are the uncinate fasciculus, the arcuate fasciculus, and the white matter connection between the hippocampus and the medial prefrontal cortex. First, the uncinate fasciculus connects the anterior temporal lobe with the medial and lateral orbitofrontal cortex (Catani & Thiebaut de Schotten, 2008). It is considered to play a major role in memory function (Mabbott et al., 2009). Several DTI studies have found an association between the uncinate fasciculus and memory performance. Christidi et al. (2017) examined the clinical relevance between white matter microstructural changes and neurophysiological measures of episodic memory and reported a relationship between episodic memory performance and the integrity of the uncinate fasciculus. Furthermore, Von Der Heide, Skipper, Klobusicky, and Olson (2013) concluded in their review that episodic memory, language, and social-emotional processing are associated with the uncinate fasciculus. Second, the arcuate fasciculus connects the perisylvian cortex of the frontal, parietal, and temporal lobes (Catani & Thiebaut de Schotten, 2008). These brain regions are known for their roles in various brain functions, such as memory, emotion, cognition, and language processing as well as language learning (Cai et al., 2019; Lebel & Beaulieu, 2009; Loui, Li, & Schlaug, 2011). Besides that, the arcuate fasciculus also seems to play a role in sound perception and production in musicians (Halwani, Loui, Rüber, & Schlaug, 2011). Hence, the arcuate fasciculus is involved in many brain processes, including declarative memory. Finally, the white matter connection between the hippocampus and the medial prefrontal cortex has also been implicated in memory processing. Preston and Eichenbaum (2013) describe that the pathway of the hippocampus to the medial prefrontal cortex is related to declarative memory. Both regions are namely involved in declarative memory. The hippocampus forms and retrieves specific memories and the medial prefrontal cortex seems to be mainly involved in memory retrieval and consolidation. Since these tracts are all, amongst others, involved in declarative memory, this study will focus on these three white matter tracts.

Most memory and language studies have measured white matter microstructure via diffusion tensor imaging (DTI). DTI measures the magnitude and orientation of anisotropic water diffusion in tissues (Christidi et al., 2017). DTI provides information about different properties, namely mean diffusivity, apparent diffusion coefficient, fractional anisotropy, axial diffusivity, and radial diffusivity. Mean diffusivity and apparent diffusion coefficient describe the molecular diffusion rate, fractional anisotropy describes the directional preference of diffusion, axial diffusivity describes the diffusion rate along the main axis of diffusion, and radial diffusivity describes the rate of diffusion across the main axis of diffusion, i.e. in the transverse direction (Soares, Marques, Alves, & Sousa, 2013). These DTI properties can be combined with behavioral measures, for example cognitive performance, to get an understanding of structure-function relationships (Madden et al., 2012). The current study will focus on the two most commonly used DTI properties, namely fractional anisotropy (FA) and mean diffusivity (MD) (Filley & Fields, 2016), since it is recommended to use multiple DTI properties due to the fact that multiple DTI properties enhance specificity and characterization of tissue microstructure (Alexander, Lee, Lazar, & Field, 2007). One way to investigate the trajectories of fiber tracts and connection patterns between different brain systems is tractography. This process consists of three steps, namely seeding, propagation, and termination. With seeding, the starting points from which the fiber bundles will be drawn are defined. The most common method to do this is based on defining regions of interest (ROIs), whereby one or more seeds are placed in each voxel of the ROI. During the propagation process the fibers are gradually generated starting from the seed(s). The termination of the fiber tracking procedure is based on well-defined criteria, to prevent tracts to go a certain way that is biologically not plausible. Common termination criteria are to set minimum FA thresholds (in adult brains usually 0.1-0.3) and turning angle thresholds (depending on which tract, usually 40-70°) (Soares et al., 2013). The FA threshold eliminates voxels that have a low probability of being white matter, instead of grey matter or CSF (Grieve, Williams, Paul, Clark, & Gordon, 2007). The angle threshold ensures that the maximum curvature is limited, to guarantee to track trajectories of water diffusion along axons (Dennis et al., 2015). Following tractography, tract specific white matter indices such as FA or MD can be related to the behavioral measures under investigation. Thus, DTI enables the non-invasive study of white matter microstructure in vivo (Soares et al., 2013).

Previous studies indicated that grammar learning is an important aspect of learning a second language, and that declarative memory is essential for L2 grammar learning. There are several white matter tracts, the above mentioned uncinate fasciculus, arcuate fasciculus, and the white matter connection between the hippocampus and the medial prefrontal cortex, that are related to declarative memory. However, it is not yet known how these white matter tracts are related to individual differences in L2 grammar learning. To the best of my knowledge, this will be the first study that looks into the relationship of memory-related white matter tracts and individual differences in L2 grammar learning.

In this study, two project aims will be investigated. The first aim of this project is to investigate the relationship between memory-related white matter tracts and L2 grammar learning. The second aim of this project is to investigate how L2 grammar learning is related to (a) declarative memory and (b) English (i.e. the second language of the participants) proficiency. In order to address the first aim, we will relate indices of white matter microstructure in three white matter tracts—the uncinate fasciculus, the arcuate fasciculus, and the white matter connection between the hippocampus and the medial prefrontal cortex—to individual differences in L2 grammar learning. We hypothesize that (a) a positive relationship exists between the FA of the three white matter tracts and L2 grammar learning performance. We also hypothesize that (b) a negative relationship exists between the MD of the three white matter tracts and L2 grammar learning. In order to address the second aim, we will (a) relate individual L2 grammar learning to declarative memory, measured via the Word Pairs I and II of the Wechsler Memory Scale (WMS-IV-NL). We will also (b) relate individual L2 grammar learning to English proficiency, measured by a composite score of the English Peabody Picture Vocabulary Test-IV (PPVT-IV) and the Word Fluency and Category Fluency of the English Verbal Fluency Test. The hypotheses of the second aim is that there will be (a) a positive relationship between L2 grammar learning and declarative memory and (b) a positive relationship between L2 grammar learning and English proficiency.

Method

Participants

In total, 37 students participated in the experiment. Seven participants were not able to finish the experiment due to Covid-19 measures and MRI-scanner malfunction and were therefore excluded from the analyses. One participant failed to improve on the grammar judgement task after the 5-day training and was therefore also excluded from the analyses. Finally, 29 participants (22 females) were included in data analyses. Their age varied from 16 to 30 years old (M = 22.9, *SD* = 4.1). The criteria to participate were right-handedness, Dutch as their native language, not raised bilingually, not familiar with the Icelandic language, no dyslexia or reading problems, MRI compatible (i.e. no metal plates or screws, vascular clips, active implant or pacemaker, permanent medical patch, non-removable piercings and metal splinters in or on their body), no claustrophobia, never been treated by a neurologist, never had a severe cerebral concussion or brain surgery, and no psychological or neurological disorder(s). The participants were recruited via SONA as well as distributed flyers. Participation was voluntary and participants received a monetary reward for their participation.

This study was approved by the ethical commission CMO region Arnhem-Nijmegen (CMO 2018-4561; CMO 2014/288). All participants gave informed consent after receiving information about the experiment and received an oral debriefing after they filled in a questionnaire about the experiment at the end of the study.

Materials

Behavioral tests

Participants completed a number of behavioral tests, see Table 1. The tests used for this study are the Hendriks, Bouman, Kessels, and Aldenkamp (2014) Word Pairs I and II of the Wechsler Memory Scale (WMS-IV-NL), the Dunn and Dunn (2007) English Peabody Picture Vocabulary Test-IV (PPVT-IV), and the Forbes-Mckay, Ellis, Shanks, and Venneri (2005) Letter Fluency (words starting with the letter ‘f’ and ‘a’) and Category Fluency (words in the category ‘animals’ and ‘fruit’) Task.

Table 1

Behavioral Test Battery

Behavioral Test Battery
Matrix Reasoning of the WAIS-V-NL
Peabody Picture Vocabulary Test-IV (PPVT-IV)
Peabody Picture Vocabulary Test-III-NL (PPVT-III-NL)
Test for Reception of Grammar (TROG-2)
Syntest NL
Word Fluency and Category Fluency of the Verbal Fluency Test (VFT)
Word Fluency and Category Fluency of the Word Fluency Test (WFT)
Alternating Serial Reaction Time (ASRT)
Word Pairs I and II of the WMS-IV-NL
Digit Span of the WAIS-IV-NL
Non Word Repetition test
STOP-IT task
3-min Reading Test (Drie-Minuten-Toets (DMT))

Icelandic pair-matching task

The pair-matching task is a memory game where participants had to connect word phrases and pictures with each other (Figure 1). In total, the pair-matching task consisted of 20 words, ten masculine Icelandic words and ten feminine Icelandic words, together with 20 corresponding pictures (Appendix 1). All word phrases started with the word ‘one’ (‘einn’ or ‘ein’) followed by a masculine or feminine noun (e.g. ‘api’ or ‘blussa’). Participants completed at least six rounds for this task, three rounds for all masculine word phrases and three rounds for all feminine word phrases. The participants needed to link all word phrases and pictures with each other correctly, otherwise they stayed in the same round. When the participants made a mistake, the right answer was shown on screen. The task took about five to ten minutes.

Figure 1. Example trial of the Icelandic pair-matching task

Grammar introduction training

During the grammar introduction training, participants learned to combine simple declarative Icelandic sentences. All sentences started with a subject noun (‘her’, meaning ‘here’) followed by a verb (‘er/ eru/ serdu’, meaning ‘is/ are/ you see’) and ended with one of the word phrases learned during the pair-matching task. In order to be able to combine these sentences correctly, the word phrases had to be correctly inflected according to gender, number, and case. In order to do so, six grammatical rules had to be implicitly learned: three grammatical cases, namely nominative-singular, nominative-plural, and accusative-singular, for two grammatical genders (masculine/ feminine). In total, the grammar training took about 20-30 minutes. The grammar training consisted of three blocks of 48 items and participants could reach three levels depending on their performance, to avoid floor- and ceiling effects. All participants started in level 1 and moved up one level for the next block if their accuracy was higher than 75%. In the first level, the word phrases consisted only of masculine and feminine nouns (Appendix 2). In the second level, adjectives (indicating colors) were added to only the masculine words. In the third level, these adjectives were added for both the masculine and feminine words (Appendix 3). Importantly, participants practiced only with 50% of the learned word phrases (five masculine and five feminine words) in this grammar training (Appendix 1).

Pre Grammar Judgement fMRI Task (Pre GJT)

Participants performed the Pre Grammar Judgement Task (GJT) in the MRI scanner, where they had to indicate whether Icelandic sentences were grammatically correct. For the baseline condition, the participants were shown Icelandic words and non-words (i.e. Icelandic words spelled backwards), and had to indicate whether they had previously learned the word or not. All 20 Icelandic words were used in this GJT (Appendix 1). There were four conditions containing 66 trials/ sentences each: (non-)words (the baseline), nominative-singular sentences, nominative-plural sentences, and accusative-singular sentences. In total, the task consisted of 264 trials and took about 20 minutes. For each stimulus type, 66.7% of the trials/ sentences were correct and 33.3% were incorrect. Incorrect sentences contained the same gender, but had a wrong grammatical rule. One trial of the GJT consisted of a fixation cross (varying between 500-1000 ms), an Icelandic sentence (presented for 3500 ms), and then another fixation cross (presented 1000 ms) (Figure 2).

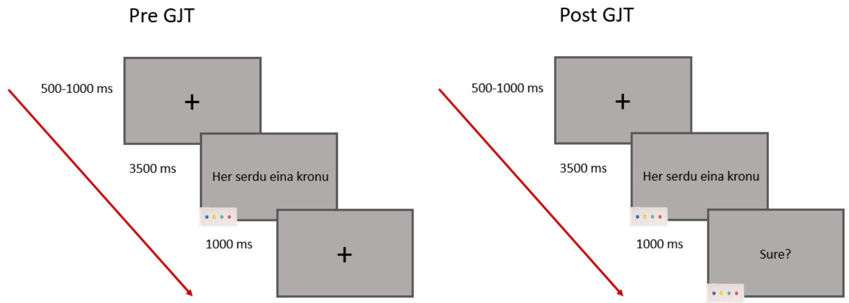

Figure 2. Example trial of the Pre (left figure) and Post (right figure) Grammar Judgement Task (GJT)

Grammar training at home

For the grammar training at home, participants had to play the grammar training every day for five consecutive days at the level reached during the grammar introduction training (Figure 3). The training took about 20-30 minutes each day. On the last day of the training, participants learned ten new Icelandic words (five masculine/ five feminine) through an identical pair-matching task as described before (Appendix 1).

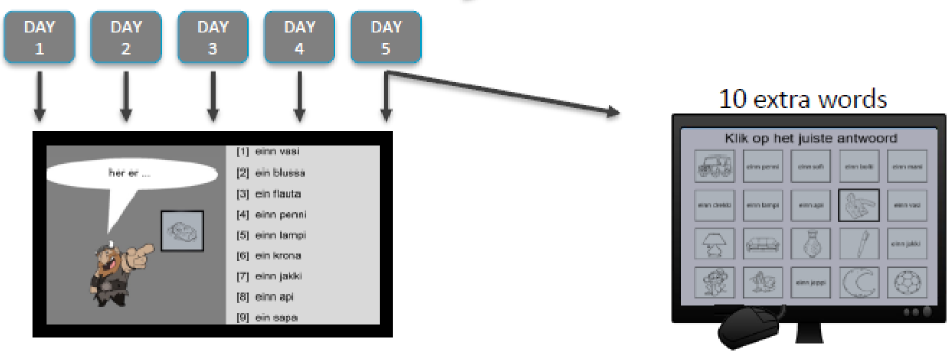

Figure 3. Example trial of the grammar training at home. On the fifth day, participants had to learn ten extra words through a similar pair-matching task.

Post Grammar Judgement fMRI Task (Post GJT)

The Post Grammar Judgement Task was very similar to the Pre GJT. One of the differences of the Post GJT was that a certainty question was added to each trial. This means that one trial of the Post GJT consisted of a fixation cross (varying between 500-1000 ms), an Icelandic sentence (presented for 3500 ms), and then the certainty question (presented 1000 ms) (Figure 2). The other difference of the Post GJT was that the ten newly learned Icelandic words were also included, which means that in total there were 30 Icelandic words in this GJT (Appendix 1). All other aspects of the Post GJT were identical to the Pre GJT.

Programs

The programs used to administer the experimental tasks used in this study were Psychopy (version 1.90) for the pair-matching task, Presentation (version 20.2) for the grammar training and GJT, and an online experiment program Gorilla (https://gorilla.sc/) for the home training.

Procedure

Participants visited the lab three times for this study (Figure 4). During the first session, participants completed a test battery consisting of different behavioral tests (Table 1). Additionally, participants learned Icelandic word phrases via an Icelandic pair-matching task (Figure 1) to prepare for the second session, which will be explained below.

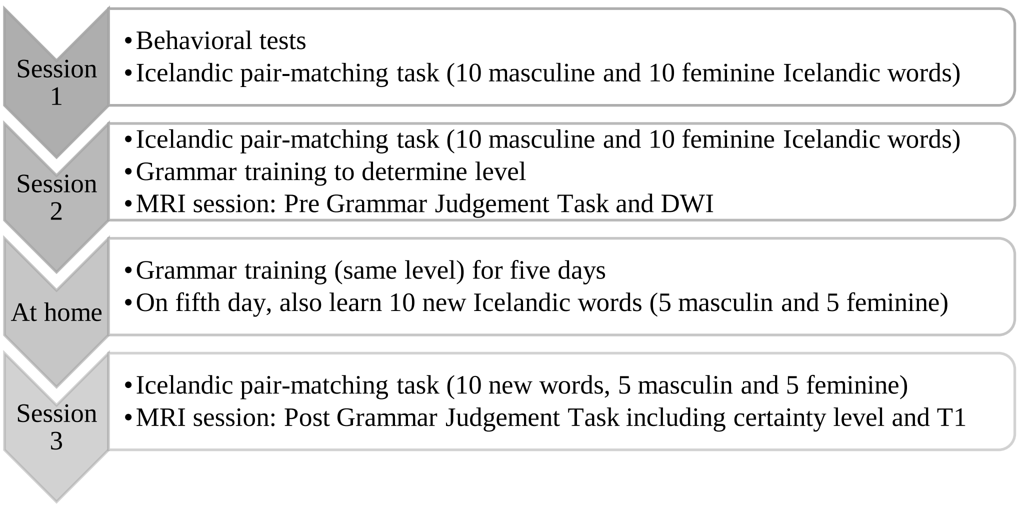

Figure 4. Testing timeline

During the second session, participants learned Icelandic grammatical rules. Icelandic was chosen because (i) some Icelandic words are cognates in Dutch, and therefore easily learned, (ii) the complexity of Icelandic grammar provided enough experimental flexibility, and (iii) every participant had no to minimal exposure to Icelandic. Before learning the grammatical rules, participants first practiced the necessary Icelandic words they had learned during session one via a pair-matching task of 20 Icelandic word phrases and pictures. Since the participants already had experience with the Icelandic pair-matching task from the first session, the Icelandic words were more consolidated before they started learning the Icelandic grammar rules. After the pair-matching task, the participants performed a grammar introduction training where they learned to combine simple declarative Icelandic sentences. After this grammar training, participants went into the MRI scanner where they performed a Pre Grammar Judgement Task (GJT) (Figure 2), where participants had to indicate whether Icelandic sentences were written correctly or incorrectly. Participants gave answers using their left index (correct) or middle finger (incorrect) on a MRI response box. During this MRI session, diffusion weighted imaging (DWI) data was collected.

Between the second and the third session, participants performed a grammar training at home for 30 minutes on five consecutive days (Figure 3). On the last day of the grammar training (training day 5), participants learned ten additional Icelandic words via the previously explained pair-matching task (Figure 1).

The third session was planned six to eight days after the second session. This session started with the participants repeating the pair-matching task from training day 5, which contained the ten new Icelandic words. The participants did not learn these words during the grammar training, so they had not encountered these Icelandic words in sentences. After the task, participants went in the MRI scanner for the second time to perform a Post Grammar Judgement Task that was similar to the Pre GJT completed during the first scan (Figure 2). The only difference was that a certainty level rating was included. The certainty question was a yes/no question asking whether the participants were confident that their given answer was correct and this question was presented after each trial. Participants gave answers using their left index (correct/ certain) or middle finger (incorrect/ uncertain) on an MRI response box. During the second MRI session, structural T1 data was also collected. At the end of the second session, all participants filled in a questionnaire to gain information about whether the participant saw patterns in the rules of the Grammar Judgement Task and if so, participants were asked to describe the patterns.

Data acquisition

Data was acquired at the Donders Institute on a 3 Tesla MAGNETOM Skyra MR scanner (Siemens AG, Healthcare Sector, Erlangen, Germany) using a product 32-channel head coil. The MRI protocol included a T1-weighted MRI scan for anatomical reference and analysis, diffusion-weighted MRI scans for probing microstructural properties and for performing fiber tractography, functional MRI data during the tasks, and resting state MRI data.

The T1-weighted scan was acquired with voxel size 1.0 mm^3, TE 2.01 ms, TR 2000 ms, TI 880 ms, FOV 256 × 256 × 192 mm, and acquisition duration 4:40 min.

The diffusion-weighted data acquisition consisted of a two-shell protocol with gradient directions that were uniformly distributed over the sphere (Caruyer, Lenglet, Sapiro, & Deriche, 2013). A custom transverse 2D PGSE pulse sequence scan was used (Setsompop et al., 2012) to acquire 6 b0 s/mm^2 volumes, 40 b1000 s/mm^2 volumes, 40 b1900 s/mm^2 volumes with voxel size 2.0 mm^3, FOV 212 mm^2, 69 slices, SMS factor 3 (interleaved), GRAPPA factor 2, phase partial Fourier Off, TE 89.60 ms, TR 2930 ms and a total acquisition duration 4:53 min. To allow for offline distortion correction of the images, 7 more b0 s/mm^2 volumes were acquired (duration 1:02 min) using the exact same sequence parameters except for the inverted k-space read-out trajectory.

Data preprocessing

The data preprocessing of this study consisted of ten steps. (1) All DWIs and MPRAGEs (Magnetization Prepared Rapid Acquisition Gradient Echo's, structural T1 images) were visually inspected for artifacts. (2) The first b0s from the anterior to posterior (AP) and posterior to anterior (PA) encoded direction were merged to create the input for TopUp. (3) We corrected for susceptibility induced distortions using FSLs TopUp (Andersson, Skare, & Ashburner, 2003; Smith et al., 2004). (4) Brain extraction (Smith, 2002) was performed, creating the brain mask needed for Eddy. (5) Eddy Cuda 9.1 was run to correct for eddy current-induced distortions and subject movements, with additional options to replace outliers, use the slice-to-volume motion model, and use the susceptibility-by-movement correction (Andersson, Graham, Drobnjak, Zhang, & Campbell, 2018; Andersson et al., 2017; Andersson, Graham, Zsoldos, & Sotiropoulos, 2016; Andersson & Sotiropoulos, 2016). (6) Eddy Quad and Eddy Squad were run (Bastiani et al., 2019). Eddy Quad was performed to generate single subject reports about potential data acquisition and pre-processing issues and to store the quality assessment indices for each subject. Eddy Squad generates study-wise reports of the Eddy Quad files and gives insight into the average movement and outliers of all participants. (7) Dtifit was run to fit the diffusion tensor model at each voxel. (8) BedpostX was run, which stands for Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques, and it is used to model crossing fibers (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007; Behrens et al., 2003). BedpostX creates all the files necessary for probabilistic tractography. (9) Registration (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001) was performed to create diffusion to standard and standard to diffusion space registrations, which are needed for probabilistic tractography. A tool to perform probabilistic tractography is ProbtrackX (Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007; Behrens et al., 2003), which produces sample streamlines in three steps. It starts from a seed, and then first draws an orientation from the voxel-wise BedpostX distributions, second takes a step in that direction, and third checks for any termination criteria. (10) The tool we used to perform probabilistic tractography is XTRACT (De Groot et al., 2013; Warrington et al., 2020), which makes use of ProbtrackX. XTRACT is a new software package for standardized and automated tractography. It can delineate white matter pathways in an automated way, using ROI-based tractography. To delineate the uncinate and arcuate fasciculus, we used seed, target and exclusion masks from XTRACT and ran XTRACT in subject's native space. It is also possible to manually select seed, target, and exclusion masks to perform tractography. This is what we did for the tract for the white matter connection between the hippocampus and the medial prefrontal cortex. Selecting the seed, target and exclusion regions for this tract was done using the same procedure as Ngo et al. (2019), who delineated the white matter connection between the hippocampus and medial prefrontal cortex using probabilistic tractography. This means that the seed was the hippocampus, the target was the medial prefrontal cortex, and the exclusion masks included the brain stem, cerebellum, occipital lobe, and parietal lobe. The midline was also included in the exclusion mask, to ensure that the modeled tract was fully lateralized. The seed and target were both obtained from the Harvard-Oxford atlas, and the exclusion regions were obtained from the MNI Structural Atlas. The three white matter tracts examined in this study are visualized in Figure 5.

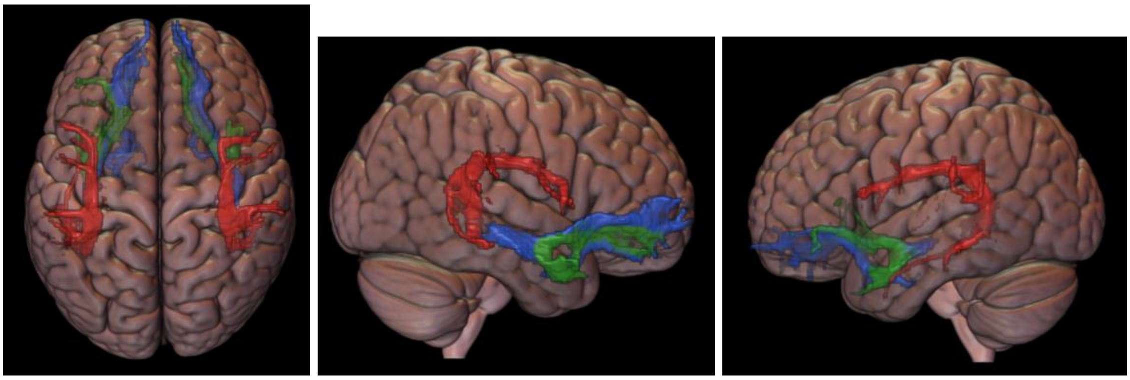

Figure 5. Superior view (left figure), right view (middle figure), and left view (right figure) of the left and right uncinate fasciculus (green), arcuate fasciculus (red), and white matter connection between the hippocampus and medial prefrontal cortex (blue) of a representative participant in MNI space.

Data analysis

Diffusion tensor imaging analysis

After tractography, the fractional anisotropy (FA) and mean diffusivity (MD) values were extracted for each tract on the left and right hemisphere separately. This way, every subject had a FA and MD value for the left and right uncinate fasciculus, arcuate fasciculus, and white matter connection between the hippocampus and the medial prefrontal cortex. A partial correlation analysis was carried out between second language (L2) grammar learning and each FA value per tract, with age, whole brain FA value, and tract volume as covariates. The same partial correlation analysis was carried out between L2 grammar learning and each MD value per tract, whilst controlling for age, whole brain MD, and tract volume. The covariates of all analyses were chosen, because we did not want these to have an effect of the relationship we are interested in, namely L2 grammar learning and each FA/ MD value per tract. All partial correlation analyses were executed in SPSS. The alpha (0.05) was corrected for multiple comparisons via the Bonferroni correction, which resulted in a corrected alpha of 0.008.

Behavioral analysis

The raw scores of the Word Pairs I and II of the WMS-IV-NL were summed to get one declarative memory total score per participant. Combining the scores to of the PPVT-IV and the Letter Fluency and Category Fluency of the Word Fluency Task gave an estimation of the participants' English (i.e. their second language) proficiency. This was done through principal component analysis (PCA), in which the two tests were used to create one composite score for English proficiency. The PCA was carried out in R, using the psych package (Revelle, 2018). To obtain L2 grammar learning, the difference in d-prime of all grammar trials between the Post GJT and Pre GJT was calculated (Figure 6), in this way the non-grammatical baseline condition was excluded from analyses. Two linear regression analyses were performed using the stats package in R: one for declarative memory and L2 grammar learning and one for English proficiency and L2 grammar learning. The declarative memory scores as well as the English proficiency scores are the independent variables (quantitative), and the L2 grammar learning is the dependent variable (quantitative).

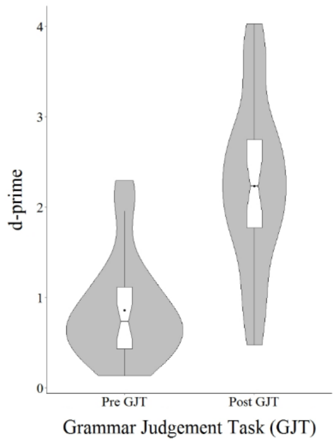

Figure 6. Violin plot of the d-prime scores of the Pre and Post Grammar Judgement Task (GJT)

Results

Behavioral results

The second language (L2) grammar learning (M = 1.43; *SD* = 0.70), English proficiency score (M = 0.04; *SD* = 1.18), and declarative memory score (M = 100.31; *SD* = 8.67) were calculated for all 29 participants (Table 2). The results of the first simple linear regression indicated a non-significant relationship ($F(1,27) = 2.42$; $p = .13$; $R^2 = .08$) between L2 grammar learning and declarative memory. The results of the second simple linear regression between L2 grammar learning and English proficiency also showed a non-significant relationship ($F(1,27) = 0.002$; $p = .97$; $R^2 < .001$).

Table 2

Descriptive Statistics of Behavioral Variables

Variable	**N**	**Mean**	**Standard Deviation**
Second language (L2) grammar learning	29	1.43	0.70
Declarative memory	24	0.41	0.40
English proficiency	29	0.04	1.18

Diffusion tensor imaging results

Table 3 presents the mean and standard deviation of the extracted fractional anisotropy (FA) and mean diffusivity (MD) values of the investigated memory-related tracts, i.e., the left and right uncinate fasciculus, arcuate fasciculus, and white matter connection between the hippocampus and the medial prefrontal cortex.

Table 3

Descriptive Statistics of Diffusion Tensor Imaging Variables

Variable		**N**	**Mean**	**SD**
Fractional Anisotropy (FA)				
Whole brain FA		29	0.48	0.01
Uncinate fasciculus	Left	29	0.40	0.02
	Right	29	0.39	0.02
Arcuate fasciculus	Left	29	0.47	0.02
	Right	29	0.47	0.02
Hippocampus to medial prefrontal cortex	Left	29	0.42	0.03
	Right	29	0.42	0.03
Mean Diffusivity (MD)				
Whole brain MD		29	0.00077	0.000019
Uncinate fasciculus	Left	29	0.00086	0.000034
	Right	29	0.00087	0.000027
Arcuate fasciculus	Left	29	0.00078	0.000025
	Right	29	0.00077	0.000024
Hippocampus to medial prefrontal cortex	Left	29	0.00089	0.000056
	Right	29	0.00087	0.000039
Tract volume				
Uncinate fasciculus	Left	29	1092.24	191.70
	Right	29	1176.21	131.01
Arcuate fasciculus	Left	29	2008.66	208.12
	Right	29	2123.72	294.07
Hippocampus to medial prefrontal cortex	Left	29	1291.62	233.99
	Right	29	1513.10	232.30

Uncinate fasciculus

We found no significant correlation between L2 grammar learning and FA of the left ($r(24) = -.02$; $p = .46$) and right ($r(24) = .33$; $p = .05$) uncinate fasciculus whilst controlling for the covariates. No significant correlation was found between MD of the left uncinate fasciculus and L2 grammar learning ($r(24) = .32$; $p = .06$), and between the MD of the right uncinate fasciculus and L2 grammar learning ($r(24) = .06$; $p = .38$) when we controlled for the covariates.

Arcuate fasciculus

For the left arcuate fasciculus no significant correlation was found between FA values and L2 grammar learning ($r(24) = .29$, $p = .07$) when controlling for the covariates. In the right hemisphere, no relation was found for the FA of the right arcuate fasciculus and L2 grammar learning ($r(24) = .35$, $p = .04$) whilst controlling for the covariates. Correlational analyses found no significant relationship between MD of the left arcuate fasciculus and L2 grammar learning ($r(24) = .06$; $p = .39$) when we controlled for the covariates. The same was found for the right hemisphere, namely no significant relationship between MD of the right arcuate fasciculus and L2 grammar learning ($r(24) = -.06$; $p = .38$) when controlling for the covariates.

Hippocampus to medial prefrontal cortex

The FA of the white matter connection between the left hippocampus and the left medial prefrontal cortex did not show a significant relationship with L2 grammar learning ($r(24) = .23$; $p = .13$) when we controlled for the covariates. The relationship between the FA of the white matter connection between the right hippocampus and the right medial prefrontal cortex and L2 grammar learning appeared to be non-significant ($r(24) = .26$; $p = .10$) whilst controlling for the covariates. There was no significant correlation found between L2 grammar learning and the MD of the white matter connection between the left ($r(24) = .09$; $p = .33$) and right ($r(24) = -.06$; $p = .38$) hippocampus and medial prefrontal cortex when controlling for the covariates.

Discussion

The main aim of this study was to examine the relationship between second language (L2) grammar learning and three bilateral memory-related white matter tracts, namely the uncinate fasciculus, the arcuate fasciculus, and the white matter connection between the hippocampus and the medial prefrontal cortex. We hypothesized that L2 grammar learning would have (a) a positive relationship with the fractional anisotropy (FA) and (b) a negative relationship with the mean diffusivity (MD) of each white matter tract. The results showed no significant correlations between FA and MD of the left and right uncinate fasciculus, arcuate fasciculus, and white matter connection between the hippocampus and the medial prefrontal cortex and the difference in d’ of the Pre and Post Grammar Judgement Task (GJT) (L2 grammar learning). The second aim of this study was to investigate how L2 grammar learning is related to (a) declarative memory and (b) English proficiency. We hypothesized that L2 grammar learning would have a positive relationship with both (a) declarative memory and (b) English proficiency. For these relationships we also did not find significant results.

Previous research has shown that the FA values of the uncinate fasciculus, the arcuate fasciculus, and the white matter connection between the hippocampus and the medial prefrontal cortex are related to declarative memory (Cai et al., 2019; Christidi et al., 2017; Preston & Eichenbaum, 2013; Von Der Heide et al., 2013). Declarative memory plays an important role in L2 grammar learning, especially in the early learning stage (Hamrick, Lum, & Ullman, 2018; Ullman, 2001). This study has been unable to demonstrate that the three memory-related white matter tracts are related to L2 grammar learning. There could be several reasons for this result. First, between-subjects design (i.e. group comparison) has less variance than within-subjects design, which increases the power of significance tests (Lane, 2013). This means that it is harder to find significance with a within-subjects design (Lane, n.d.). Therefore, this may have been one of the reasons we did not find significant results. Second, this study had a rather small sample size, which could increase the likelihood of Type II errors and thus decrease the power of this study (Deziel, 2018). In sum, the within-subject design and the small sample size of this study may have caused the non-significant results.

Additionally, this study did not detect any evidence for a relationship between L2 grammar learning and declarative memory and English proficiency, in contrast with previous studies (Golonka, 2010; Ullman, 2001). A possible explanation for not finding a relationship between L2 grammar learning and declarative memory may be due to the level of experience with the new grammar. According to the meta-analyses of Hamrick, Lum, and Ullman (2018), L2 grammar learning relies more on the declarative memory system at earlier stages of learning, but with increasing second language experience, L2 grammar learning becomes more dependent on the procedural memory system. It is feasible that the participants became familiar enough with the second language grammar of this study, by means of the grammar training, that the L2 grammar learning depended more on their procedural memory. The scores of the Post GJT namely increased compared to the Pre GJT scores. One study particularly shows a shift from declarative memory to procedural memory after one to three weeks without in-between exposure (Hamrick, 2015). In the current study, we tested participants on their grammatical knowledge after six to eight days after five days of grammar training. Hence, there is a possibility that the L2 grammar learning shifted from declarative to procedural memory within this training, whereby the relationship between declarative memory and L2 grammar learning did not come to light.

The reason for not finding a significant relationship between L2 grammar learning and English proficiency could be related to how we defined English language proficiency. We estimated English proficiency using a composite score of the English Peabody Picture Vocabulary Test-IV and the Letter Fluency and Category Fluency of the Word Fluency Test, both language tests related to vocabulary (Unruh & McKellar, 2017). One study that did find a positive relationship between second and third language grammar proficiency (Moghtadi, Koosha, and Lotfi, 2014) used a grammatical proficiency test to measure the second (English) and third (Persian) language grammatical proficiency of their participants. Our study originally planned to include a different L2 grammatical test, namely the Bishop (2003) Test for Reception of Grammar (TROG-2). However, the results of this test did not meet up to the assumptions of the PCA, which led to the decision to leave this test out of the English proficiency score. However, future studies could use an English grammatical proficiency test together with the two English vocabulary-related tests to get an English language proficiency score that consists of both L2 vocabulary-related and grammatical tests.

This study had several limitations. First, the sample size was smaller than planned due to the Covid-19 restriction and the MRI-scanner being out of service for several weeks. The power dropped due to the small sample size, which likely contributed to the non-significant findings. Second, the learning environment for the grammar training at home was different for each participant. Ideally, the grammar training should have taken place in a lab environment to ensure that each participant has an identical environment. However, since this study already required the participants to come to the lab often, we chose to let the participants carry out the grammar training online to make dropout less likely. Third, as stated above, the composite score for English proficiency consisted only of two tests related to vocabulary, instead of also a grammatical test. Future research should also include a L2 grammatical test to make sure the English proficiency score consists of both L2 vocabulary-related and grammatical tests.

Altogether, the findings of this study did not show support of any relationship between the uncinate fasciculus, arcuate fasciculus, and white matter connection between the hippocampus and the medial prefrontal cortex and L2 grammar learning. In our sample, memory-related white matter tracts are not associated with L2 grammar learning. Still, more research is necessary to fully support or reject this finding. Overall, this study has shed more light on the relationship between several memory-related white matter tracts and L2 grammar learning, and therefore contributes to improve our knowledge in this area.

References

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion Tensor Imaging of the Brain. *Neurotherapeutics*, *4*(3), 316–329. https://doi.org/10.1016/j.nurt.2007.05.011

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., & Campbell, J. (2018). Susceptibility-induced distortion that varies due to motion: Correction in diffusion MR without acquiring additional data. *NeuroImage*, *171*, 277–295. https://doi.org/10.1016/j.neuroimage.2017.12.040

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., Filippini, N., & Bastiani, M. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion MR images: Within volume movement. *NeuroImage*, *152*, 450–466. https://doi.org/10.1016/j.neuroimage.2017.02.085

Andersson, J. L. R., Graham, M. S., Zsoldos, E., & Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *NeuroImage*, *141*, 556–572. https://doi.org/10.1016/j.neuroimage.2016.06.058

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., … 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

Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, *34*(1), 144–155. https://doi.org/10.1016/j.neuroimage.2006.09.018

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., … Smith, S. M. (2003). Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging. *Magnetic Resonance in Medicine*, *50*(5), 1077–1088. https://doi.org/10.1002/mrm.10609

Bishop, D. (2003). Test for Reception of Grammar (TROG-2) [Measurement instrument]. Retrieved from https://www.pearsonclinical.co.uk/Psychology/ChildCognitionNeuropsychologyandLanguage/ChildLanguage/TestforReceptionofGrammar(TROG-2)/TestforReceptionofGrammar(TROG-2).aspx

Cai, W., Zhao, M., Liu, J., Liu, B., Yu, D., & Yuan, K. (2019). Right arcuate fasciculus and superior longitudinal fasciculus abnormalities in primary insomnia. *Brain Imaging and Behavior*, *13*(6), 1746–1755. https://doi.org/10.1007/s11682-019-00160-1

Caruyer, E., Lenglet, C., Sapiro, G., & Deriche, R. (2013). Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance in Medicine*, *69*(6), 1534–1540. https://doi.org/10.1002/mrm.24736

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132. https://doi.org/10.1016/j.cortex.2008.05.004

Christidi, F., Karavasilis, E., Zalonis, I., Ferentinos, P., Giavri, Z., Wilde, E. A., … Evdokimidis, I. (2017). Memory-related white matter tract integrity in amyotrophic lateral sclerosis: an advanced neuroimaging and neuropsychological study. *Neurobiology* *of Aging*, *49*, 69–78. https://doi.org/10.1016/j.neurobiolaging.2016.09.014

De Groot, M., Vernooij, M. W., Klein, S., Ikram, M. A., Vos, F. M., Smith, S. M., … Andersson, J. L. R. (2013). Improving alignment in Tract-based spatial statistics: Evaluation and optimization of image registration. *NeuroImage*, *76*, 400–411. https://doi.org/10.1016/j.neuroimage.2013.03.015

Dennis, E. L., Jin, Y., Kernan, C., Babikian, T., Mink, R., Babbitt, C., … Thompson, P. M. (2015). White matter integrity in traumatic brain injury: Effects of permissible fiber turning angle. *Proceedings - International Symposium on Biomedical Imaging*, *2015–July*, 930–933. https://doi.org/10.1109/ISBI.2015.7164023

Deziel, C. (2018). The Effects of a Small Sample Size Limitation. Retrieved from https://sciencing.com/effects-small-sample-size-limitation-8545371.html

Dunn, L. M., & Dunn, D. M. (2007). Peabody Picture Vocabulary Test | Fourth Edition [Measurement instrument]. Retrieved from https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/Brief/Peabody-Picture-Vocabulary-Test-%7C-Fourth-Edition/p/100000501.html?tab=pricing-ordering

Eurostat. (2019). Foreign language skills statistics - Statistics Explained. In *Statistics Explained*. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Category:Tourism_glossary

Filley, C. M., & Fields, R. D. (2016, November 1). White matter and cognition: Making the connection. *Journal of Neurophysiology*, Vol. 116, pp. 2093–2104. https://doi.org/10.1152/jn.00221.2016

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*(4), 527–552. https://doi.org/10.1017/S0272263101004041

Flöel, A., de Vries, M. H., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2009.05.046

Forbes-Mckay, K. E., Ellis, A. W., Shanks, M. F., & Venneri, A. (2005). The age of acquisition of words produced in a semantic fluency task can reliably differentiate normal from pathological age related cognitive decline. *Neuropsychologia*, *43*(11), 1625–1632. https://doi.org/10.1016/j.neuropsychologia.2005.01.008

Golonka, E. (2010). The Influence of Prior Language Learning Experiences on Learning of Unrelated or Distantly Related Languages. *Russian Language Journal*, *60*, 95–121.

Grieve, S. M., Williams, L. M., Paul, R. H., Clark, C. R., & Gordon, E. (2007). Cognitive aging, executive function, and fractional anisotropy: A diffusion tensor MR imaging study. *American Journal of Neuroradiology*, *28*(2), 226–235.

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, *2*(JUL). https://doi.org/10.3389/fpsyg.2011.00156

Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9–15. https://doi.org/10.1016/j.lindif.2015.10.003

Hamrick, P., Lum, J. A. G., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(7), 1487–1492. https://doi.org/10.1073/pnas.1713975115

He, H. (2013). On FL learners' Individual differences in grammar learning and their grammatical competence training. *Theory and Practice in Language Studies*, *3*(8), 1369–1374. https://doi.org/10.4304/tpls.3.8.1369-1374

Hendriks, M. P. H., Bouman, Z., Kessels, R. P. C., & Aldenkamp, A. P. (2014). Wechsler Memory Scale IV-NL [Measurement instrument]. Retrieved from https://www.pearsonclinical.nl/wms-iv-nl-wechsler-memory-scale

Huber, E., Donnelly, P. M., Rokem, A., & Yeatman, J. D. (2018). Rapid and widespread white matter plasticity during an intensive reading intervention. *Nature Communications*, *9*(1), 2260. https://doi.org/10.1038/s41467-018-04627-5

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage, 17*(2), 825–841. https://doi.org/10.1016/S1053-8119(02)91132-8

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. https://doi.org/10.1016/S1361-8415(01)00036-6

Johansen-Berg, H. (2010). Behavioural relevance of variation in white matter microstructure. *Current Opinion in Neurology*, 1. https://doi.org/10.1097/WCO.0b013e32833b7631

Lahmann, C., Steinkrauss, R., & Schmid, M. S. (2016). Factors Affecting Grammatical and Lexical Complexity of Long-Term L2 Speakers' Oral Proficiency. *Language Learning*, *66*(2), 354–385. https://doi.org/10.1111/lang.12151

Lane, D. M. (n.d.). Online Statistics Education: A Multimedia Course of Study. Retrieved from http://onlinestatbook.com/

Lane, D. M. (2013). HyperStat Online Statistics Textbook. Retrieved from http://davidmlane.com/hyperstat/index.html

Lebel, C., & Beaulieu, C. (2009). Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Human Brain Mapping, 30*(11), 3563–3573. https://doi.org/10.1002/hbm.20779

Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage*, *55*(2), 500–507. https://doi.org/10.1016/j.neuroimage.2010.12.022

Mabbott, D. J., Rovet, J., Noseworthy, M. D., Smith, M. Lou, & Rockel, C. (2009). The relations between white matter and declarative memory in older children and adolescents. *Brain Research*, *1294*, 80–90. https://doi.org/10.1016/j.brainres.2009.07.046

Madden, D. J., Bennett, I. J., Burzynska, A., Potter, G. G., Chen, N. kuei, & Song, A. W. (2012, March 1). Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, Vol. 1822, pp. 386–400. https://doi.org/10.1016/j.bbadis.2011.08.003

Moghtadi, L., Koosha, M., & Lotfi, A. R. (2014). Second language grammatical proficiency and third language acquisition. *International Education Studies*, *7*(11), 19–27. https://doi.org/10.5539/ies.v7n11p19

Ngo, C. T., Alm, K. H., Metoki, A., Hampton, W., Riggins, T., Newcombe, N. S., & Olson, I. R. (2019). White matter structural connectivity and episodic memory in early childhood. *Developmental Cognitive Neuroscience*, *28*, 41–53. https://doi.org/10.1016/j.dcn.2017.11.001

Preston, A. R., & Eichenbaum, H. (2013, September 9). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, Vol. 23, pp. R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

Revelle, W. (2018) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.12.

Roediger, H. L., Zaromb, F. M., & Lin, W. (2017). A Typology of Memory Terms. In *Learning and Memory: A Comprehensive Reference* (pp. 7–19).

https://doi.org/10.1016/B978-0-12-809324-5.21003-1

Setsompop, K., Cohen-Adad, J., Gagoski, B. A., Raij, T., Yendiki, A., Keil, B., … Wald, L. L. (2012). Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *NeuroImage*, *63*(1), 569–580. https://doi.org/10.1016/j.neuroimage.2012.06.033

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2004.07.051

Soares, J. M., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neuroscience*, *7*(7 MAR), 1–14. https://doi.org/10.3389/fnins.2013.00031

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(1), 105–122. https://doi.org/10.1017/S1366728901000220

Ullman, M. T., & Lovelett, J. T. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, *34*(1), 39–65. https://doi.org/10.1177/0267658316675195

Unruh, S., & McKellar, N. A. (2017). Reading Fluency and Vocabulary. In *Assessment and Intervention for English Language Learners* (pp. 81–96). https://doi.org/10.1007/978-3-319-52645-4_6

Unsworth, S., Persson, L., Prins, T., & De Bot, K. (2015). An Investigation of Factors Affecting Early Foreign Language Learning in the Netherlands. *Applied Linguistics*, *36*(5), 527–548. https://doi.org/10.1093/applin/amt052

Von Der Heide, R. J., Skipper, L. M., Klobusicky, E., & Olson, I. R. (2013). Dissecting the uncinate fasciculus: disorders, controversies and a hypothesis. *Brain : A Journal of Neurology*, *136*(Pt 6), 1692–1707. https://doi.org/10.1093/brain/awt094

Warrington, S., Bryant, K. L., Khrapitchev, A. A., Sallet, J., Charquero-Ballester, M., Douaud, G., … Sotiropoulos, S. N. (2020). XTRACT - Standardised protocols for automated tractography in the human and macaque brain. *NeuroImage*, *217*, 116923. https://doi.org/10.1016/j.neuroimage.2020.116923

Zidane, R. (2016). The role of memory in foreign language learning and achievement. *Humanitas*, *4*(7), 499–516. https://doi.org/http://dx.doi.org/10.20304/husbd.84543

Appendices

Appendix 1. Stimuli list of Icelandic words

	Masculine words	Feminine words	Pair-matching task session 1	Grammar introduction training	Pre GJT	Grammar training at home	Pair-matching task day 5	Post GJT
1	einn api	ein blussa	x	x	x	x		x
2	einn jakki	ein flauta	x	x	x	x		x
3	einn lampi	ein krona	x	x	x	x		x
4	einn penni	ein panna	x	x	x	x		x
5	einn vasi	ein sapa	x	x	x	x		x
6	einn sofi	ein terta	x		x	x		x
7	einn bolti	ein taska	x		x	x		x
8	einn dreki	ein flaska	x		x	x		x
9	einn jeppi	ein pumpa	x		x	x		x
10	einn mani	ein klukka	x		x	x		x
11	einn bursti	ein kirkja					x	x
12	einn falki	ein pera					x	x
13	einn pakki	ein tromma					x	x
14	einn safi	ein stjarna					x	x
15	einn hani	ein kista					x	x

Appendix 2. Stimuli list of Icelandic sentences without adjectives (masculine and feminine)

	Masculine words		
	Nominative-singular	Nominative-plural	Accusative-singular
1	her er einn api	her eru tveir apar	her serdu einn apa
2	her er einn jakki	her eru tveir jakkar	her serdu einn jakka
3	her er einn lampi	her eru tveir lampar	her serdu einn lampa
4	her er einn penni	her eru tveir pennar	her serdu einn penna
5	her er einn vasi	her eru tveir vasar	her serdu einn vasa
6	her er einn sofi	her eru tveir sofar	her serdu einn sofa
7	her er einn bolti	her eru tveir boltar	her serdu einn bolta
8	her er einn dreki	her eru tveir drekar	her serdu einn dreka
9	her er einn jeppi	her eru tveir jeppar	her serdu einn jeppa
10	her er einn mani	her eru tveir manar	her serdu einn mana
11	her er einn bursti	her eru tveir burstar	her serdu einn bursta
12	her er einn falki	her eru tveir falkar	her serdu einn falka
13	her er einn pakki	her eru tveir pakkar	her serdu einn pakka
14	her er einn safi	her eru tveir safar	her serdu einn safa
15	her er einn hani	her eru tveir hanar	her serdu einn hana

	Feminine words		
	Nominative-singular	Nominative-plural	Accusative-singular
1	her er ein blussa	her eru tvaer blussur	her serdu eina blussu
2	her er ein flauta	her eru tvaer flautur	her serdu eina flautu
3	her er ein krona	her eru tvaer kronur	her serdu eina kronu
4	her er ein panna	her eru tvaer pannur	her serdu eina pannu
5	her er ein sapa	her eru tvaer sapur	her serdu eina sapu
6	her er ein terta	her eru tvaer tertur	her serdu eina tertu
7	her er ein taska	her eru tvaer taskur	her serdu eina tasku
8	her er ein flaska	her eru tvaer flaskur	her serdu eina flasku
9	her er ein pumpa	her eru tvaer pumpur	her serdu eina pumpu
10	her er ein klukka	her eru tvaer klukkur	her serdu eina klukku
11	her er ein kirkja	her eru tvaer kirkjur	her serdu eina kirkju
12	her er ein pera	her eru tvaer perur	her serdu eina peru
13	her er ein tromma	her eru tvaer trommur	her serdu eina trommu
14	her er ein stjarna	her eru tvaer stjarnur	her serdu eina stjarnu
15	her er ein kista	her eru tvaer kistur	her serdu eina kistu

Appendix 3. Stimuli list of Icelandic sentences with adjectives (masculine and feminine). The examples are with the color yellow (gulur). Green, white and black (graent, hvitur, and svartur respectively) are the other three adjectives.

	Masculine words		
	Nominative-singular	Nominative-plural	Accusative-singular
1	her er einn gulur api	her eru tveir gulir apar	her serdu einn gulan apa
2	her er einn gulur jakki	her eru tveir gulir jakkar	her serdu einn gulan jakka
3	her er einn gulur lampi	her eru tveir gulir lampar	her serdu einn gulan lampa
4	her er einn gulur penni	her eru tveir gulir pennar	her serdu einn gulan penna
5	her er einn gulur vasi	her eru tveir gulir vasar	her serdu einn gulan vasa
6	her er einn gulur sofi	her eru tveir gulir sofar	her serdu einn gulan sofa
7	her er einn gulur bolti	her eru tveir gulir boltar	her serdu einn gulan bolta
8	her er einn gulur dreki	her eru tveir gulir drekar	her serdu einn gulan dreka
9	her er einn gulur jeppi	her eru tveir gulir jeppar	her serdu einn gulan jeppa
10	her er einn gulur mani	her eru tveir gulir manar	her serdu einn gulan mana
11	her er einn gulur bursti	her eru tveir gulir burstar	her serdu einn gulan bursta
12	her er einn gulur falki	her eru tveir gulir falkar	her serdu einn gulan falka
13	her er einn gulur pakki	her eru tveir gulir pakkar	her serdu einn gulan pakka
14	her er einn gulur safi	her eru tveir gulir safar	her serdu einn gulan safa
15	her er einn gulur hani	her eru tveir gulir hanar	her serdu einn gulan hana

	Feminine words		
	Nominative-singular	Nominative-plural	Accusative-singular
1	her er ein gul blussa	her eru tvaer gular blussur	her serdu eina gula blussu
2	her er ein gul flauta	her eru tvaer gular flautur	her serdu eina gula flautu
3	her er ein gul krona	her eru tvaer gular kronur	her serdu eina gula kronu
4	her er ein gul panna	her eru tvaer gular pannur	her serdu eina gula pannu
5	her er ein gul sapa	her eru tvaer gular sapur	her serdu eina gula sapu
6	her er ein gul terta	her eru tvaer gular tertur	her serdu eina gula tertu
7	her er ein gul taska	her eru tvaer gular taskur	her serdu eina gula tasku
8	her er ein gul flaska	her eru tvaer gular flaskur	her serdu eina gula flasku
9	her er ein gul pumpa	her eru tvaer gular pumpur	her serdu eina gula pumpu
10	her er ein gul klukka	her eru tvaer gular klukkur	her serdu eina gula klukku
11	her er ein gul kirkja	her eru tvaer gular kirkjur	her serdu eina gula kirkju
12	her er ein gul pera	her eru tvaer gular perur	her serdu eina gula peru

13	her er ein gul tromma	her eru tvaer gular trommur	her serdu eina gula trommu
14	her er ein gul stjarna	her eru tvaer gular stjarnur	her serdu eina gula stjarnu
15	her er ein gul kista	her eru tvaer gular kistur	her serdu eina gula kistu