博士論文

Functional anatomy of the dorsal and ventral language-related pathways in second language acquisition

(第二言語習得における背側および腹側言語関連経路の機能解剖学)

山本香弥子

Table of contents

List of tables ····· 4

List of figures ····· 5

List of abbreviations ····· 6

Chapter 1. General introduction ····· 8

Chapter 2. The dorsal rather than ventral pathway better reflects individual syntactic abilities in second language ····· 13

2.1. Introduction ····· 14

2.2. Materials and methods ····· 18

2.2.1. Experiment I ····· 18

2.2.1.1. Participants ····· 18

2.2.1.2. Stimuli and tasks ····· 20

2.2.1.3. MR image acquisition ····· 22

2.2.1.4. Diffusion MRI data analyses ····· 22

2.2.1.5. Identification of the Arcuate and IFOF ····· 23

2.2.1.6. ROI selection to minimize individual variances in the fiber tract's thickness ··· 26

2.2.1.7. FA analyses ····· 29

2.2.2. Experiment II ····· 29

2.2.3. Effects of different imaging parameters on FA ····· 30

2.3. Results ····· 30

2.3.1. Experiment I ····· 30

2.3.1.1. Behavioral data ····· 30

2.3.1.2. Structural properties of the Arcuate and IFOF ····· 33

2.3.1.3. FA of the left Arcuate as an indicator of individual performances for the Syn Task ····· 34

2.3.2. Experiment II ····· 36
2.3.2.1. Behavioral correlations within twin pairs ····· 36
2.3.2.2. Structural correlations within twin pairs ····· 37
2.3.3. Effects of the imaging parameters on FA ····· 39
2.4. Discussion ····· 41
Chapter 3. Differential signatures of second language syntactic performance and age on the structural properties of the left dorsal pathway ····· 49
3.1. Introduction ····· 50
3.2. Materials and methods ····· 53
3.2.1. Participants ····· 53
3.2.2. Identification of the Arcuate and IFOF ····· 55
3.2.3. ROI selection ····· 56
3.2.4. FA analyses ····· 57
3.3. Results ····· 58
3.3.1. Behavioral data ····· 58
3.3.2. Group differences along the tract for the Arcuate ····· 60
3.3.3. Distinct group differences in the structural properties of the left Arcuate ····· 63
3.3.4. FA in the left Arcuate was selectively correlated with the accuracy of Syn ····· 65
3.4. Discussion ····· 67
Chapter 4. General discussion ····· 71
Conclusions ····· 75
Acknowledgements ····· 77
References ····· 78

List of tables

Table 2.1. Demographic details of the participants in each experiment. ····· 20

Table 2.2. Behavioral data of the participants. ····· 31

Table 2.3. Structural properties of the Arcuate and IFOF. ····· 33

Table 3.1. Demographic details of the participants in each group. ····· 55

List of figures

Figure 1.1. The dorsal and ventral language-related pathways in the human brain. ····· 9

Figure 1.2. Diffusion anisotropy represented by fractional anisotropy (FA). ····· 10

Figure 2.1. The Arcuate and inferior fronto-occipital fasciculus (IFOF) reconstructed by diffusion magnetic resonance imaging (MRI) in a typical participant. ····· 25

Figure 2.2. Region of interest (ROI) selection to minimize individual variances in the fiber tract's thickness. ····· 28

Figure 2.3. Independence of the accuracy of the Syn and Spe tasks. ····· 32

Figure 2.4. FA of the left Arcuate as an indicator of individual performances for the Syn task. ····· 35

Figure 2.5. Behavioral correlations within twin pairs. ····· 37

Figure 2.6. The structural correlation of the left Arcuate within twin pairs. ····· 38

Figure 2.7. The consistency of FA for different imaging parameters. ····· 40

Figure 3.1. Behavioral data of Syn and Spe tasks for the three groups. ····· 59

Figure 3.2. The Arcuate and IFOF reconstructed with diffusion MRI. ····· 61

Figure 3.3. Profiles of the thickness/FA of the dorsal and ventral pathways in each group. · 62

Figure 3.4. Group differences in the structural properties of the left Arcuate. ····· 64

Figure 3.5. FA in the left Arcuate was selectively correlated with the accuracy of Syn. ····· 66

List of abbreviations

Abbreviation	Meaning
A	anterior
AC-PC	anterior to posterior commissure
AOE	age of first exposure
Arcuate	arcuate fasciculus
cocor	comparing correlations
DOE	duration of exposure
EC	external capsule
EmC	extreme capsule
ENIGMA	Enhancing Neuroimaging Genetics through Meta-Analysis
F3t	triangular part of the left inferior frontal gyrus
FA	fractional anisotropy
FDT	FMRIB's Diffusion Toolbox
FLIRT	FMRIB's Linear Image Registration Tool
FMRIB	Functional MRI of the Brain
FNIRT	FMRIB's Nonlinear Image Registration Tool
FOV	field of view
FSL	FMRIB's Software Library
HCP	Human Connectome Project
I	inferior
IFG	inferior frontal gyrus
IFOF	inferior fronto-occipital fasciculus
ILF	inferior longitudinal fasciculus
L	left
L1	first language
L2	second language
LQ	laterality quotient
MdLF	middle longitudinal fasciculus
MEG	magnetoencephalography
MEXT	Ministry of Education, Culture, Sports, Science and Technology
MNI	Montreal Neurological Institute

Abbreviation	Meaning
MR(I)	magnetic resonance (imaging)
N	the number of participants
ODF	orientation distribution functions
P	posterior
ppcor	partial and semi-partial correlation
QBOOT	*q*-ball orientation distribution functions and residual bootstrap
R	right
rANOVA	repeated-measures analysis of variance
ROI	region of interest
RTs	reaction times
S	superior
SD	standard deviation
SLF	superior longitudinal fasciculus
Spe	spelling error-detection
Syn	syntactic error-detection
TBSS	tract-based spatial statistics
TE	echo time
TR	repetition time
UF	uncinate fasciculus
VBM	voxel-based morphometry
VWFA	visual word form area

Chapter 1. General introduction

In the human brain, language functions are supported by distributed regions especially in the left hemisphere. It has been widely accepted that some of these cortical regions are mutually connected by the dorsal and ventral pathways, i.e., language-related pathways (Figure 1.1) (Hickok and Poeppel, 2007; Friederici, 2011).

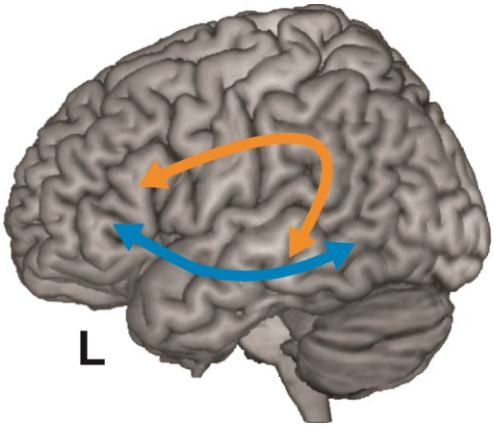

Figure 1.1. The dorsal and ventral language-related pathways in the human brain. Cortical regions supporting language functions are anatomically connected by the dorsal (orange) and ventral (blue) pathways.

The recent development of diffusion magnetic resonance imaging (MRI) techniques has opened an avenue for the structural studies of fiber bundles in a non-invasive manner, by enabling, for example, the assessment of each bundle's fractional anisotropy (FA)—the degree of fiber organization and/or myelination. FA is calculated as follows using the eigenvalues (mm^2/s) of a three-dimensional ellipsoid that represents water diffusion within each voxel (Figure 1.2A).

$$\mathrm{FA} = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{{\lambda_1}^2 + {\lambda_2}^2 + {\lambda_3}^2}} \quad (\bar{\lambda} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3})$$

FA is scaled from 0 to 1; FA is 0 when the diffusion is isotropic, and FA is 1 when the diffusion is completely anisotropic (Figure 1.2B). Higher FA indicates higher diffusion anisotropy, which is associated with the existence of fibers that hamper water diffusion. The

longest axis of the ellipsoid is thus assumed to indicate the fiber orientation. The bundle's volume or thickness is another structural property independent of FA, since it is possible that the tract becomes thicker even when the tract has lower FA. Cellular mechanisms include fiber organization and myelin formation/remodeling for increases in FA, as well as astrocyte growth and angiogenesis for increases in the volume.

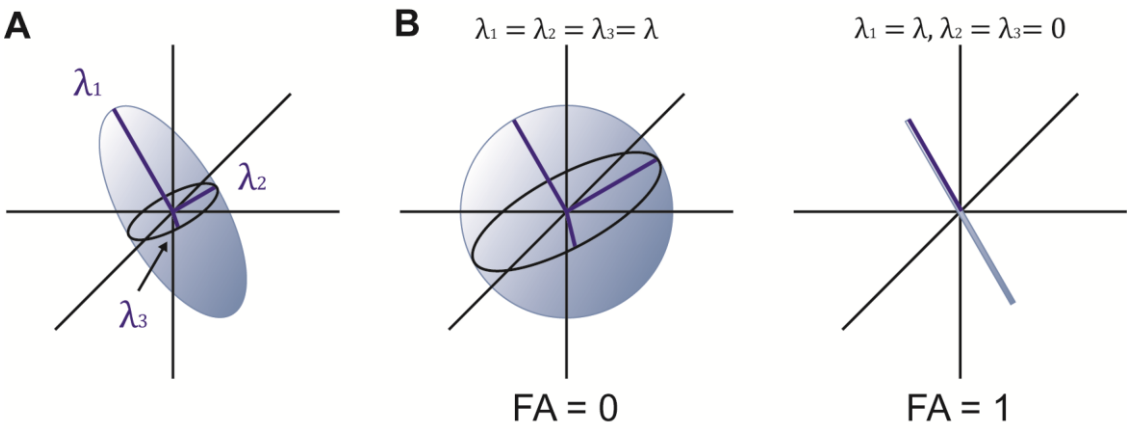

Figure 1.2. Diffusion anisotropy represented by fractional anisotropy (FA). (A) Water diffusion estimated by a three-dimensional ellipsoid model. (B) Examples where FA takes lowest (FA = 0; left) and highest (FA = 1; right) values.

FA has been reported to increase in white matter underlying the intraparietal sulcus after a 6-week training of a visuo-motor skill (Scholz et al., 2009), as well as in the genu of the corpus callosum during a 9-month training of a second language (L2, a secondary language acquired mostly at school) (Schlegel et al., 2012). These studies have induced researchers to examine a correlation between FA and performance metrics. In order to reveal the functional roles of the language-related pathways, it would be important to separate syntactic or semantic abilities from other general skills, and to further identify which structural properties of dorsal and/or ventral pathways actually reflect individual differences in these abilities. Here, I examined individual syntactic abilities in L2 among Japanese high-school students, who started to acquire English at the age of 12 or 13, thereby exploring how

individual differences in syntactic acquisition were related to the microscopic properties of the dorsal and ventral language-related pathways.

Second language acquisition shows considerably large individual differences, especially when the L2 is acquired in adulthood. Even among people with similar L2 experiences (e.g., taking the same classes or lessons on a foreign language), some improve their L2 performances in a relatively short period of time, while others do not improve as well. This is in marked contrast to first language (L1, a native language) acquisition, in which linguistic abilities are similar among individuals despite highly variable experiences. However, individual differences in L1 do emerge when effortful language use is imposed. By using verb-, rhyme-, and opposite word generation tasks in L1 for nine-year-old children and adult participants, previous studies have reported that distinct regions in the left frontal and occipital cortices showed age-related or performance-related activations (Schlaggar et al., 2002; Brown et al., 2005). These studies have indicated the importance of employing subgroups, in which either age or performance was matched, thereby fixing one of the two factors in group comparisons. Individual differences would also be revealed when people learn to read or write even with their own languages, which requires educational training of specific skills. A previous study has reported that the structural property of a white matter pathway connecting the left temporal and parietal language areas may have plasticity associated with literacy experience even for adults, by comparing three groups, i.e., illiterates, ex-illiterates who learned to read during adulthood, and literates (Thiebaut de Schotten et al., 2014). Long and intensive experiences are required also in L2 acquisition, which may be supported by the structural plasticity of language-related pathways. To elucidate such neuroplasticity, it is of interest to focus on the transitional period of adolescence, which occurs after the sensitive period of language acquisition (≤12 years old). Indeed, this period has been suggested to involve certain plasticity towards maturation (Fuhrmann et al., 2015). The adolescent brain would thus be influenced by factors depending on the participants' ages

(e.g., biological maturation), as well as performances in L2. More specifically, the independent factors of age and performances may be reflected in the plasticity of different structural properties in major language-related pathways.

In this dissertation, I aimed to reveal which of the dorsal and ventral language-related pathways was more critical to the individual syntactic abilities in L2. I then recruited monozygotic twins and examined to what extent their L2 abilities and their structural properties were similar. Further, by recruiting more adolescent students, I examined individual differences in L2 performances controlling the duration of experience to reveal the differential signatures of performances and age on the plasticity of structural properties in the major language-related pathways. This study is in the area of functional anatomy, which elucidates brain functions based on the cortical structures.

Chapter 2.

The dorsal rather than ventral pathway better reflects individual syntactic abilities in second language

2.1. Introduction

When tracking pathways in an individual brain with diffusion MRI, we should assess the structural properties of a certain tract with caution, because in some studies not all of the *right* arcuate fasciculus (Arcuate), which overlaps with the superior longitudinal fasciculus (SLF), were reconstructed (Catani et al., 2007; Lebel and Beaulieu, 2009; Thiebaut de Schotten et al., 2011; López-Barroso et al., 2013). Possible causes of such uncertainty may be the limitations of deterministic tracking algorithms (Yeatman et al., 2011), and/or other methods for tracking bundles—for example, starting points (i.e., seeds) were drawn manually in many studies, making the strict inclusion/exclusion of certain fibers difficult. The use of probabilistic tracking algorithms, as well as a reliable and objective method for seed definition, is thus necessary. Another difficulty in examining structural properties of individual brains is that macroscopic differences may depend on a number of structural factors, such as the degrees of arching and branching, as well as individual variations in the shape and size of tracts, especially near the gray matter (Catani et al., 2005; Anwander et al., 2007). One approach for focusing on areas of small inter-subject variability is to select regions of interest (ROIs) that exclude potentially variable branching regions (Tsang et al., 2009; Yeatman et al., 2011). Although there is another strategy of Tract-Based Spatial Statistics (TBSS), in which voxels in the center of fiber bundles with the highest FA (i.e., the "mean FA skeleton") are examined (Smith et al., 2006), it has been argued that the mean FA skeleton is a very limited portion at the center (Van Hecke et al., 2010; Bach et al., 2014). Selecting ROIs has an advantage over TBSS, such that the thickness (or volume), i.e., another structural property, can be obtained at a full-width cross-section of a fiber bundle. Here we employed this ROI approach, and tried to improve the method of determining the size and position of ROIs objectively.

In the present diffusion MRI study with probabilistic tracking, we developed the following semi-automatic methods, and minimized the manual procedures. Seeds were

defined in a semi-automatic and consistent manner, and the tracts were successfully reconstructed in both hemispheres for every participant. We then selected a ROI in each tract at the portion with the most uniform *thickness*, thereby excluding the branching or curved portions. The ROI size for each pathway was also optimized to minimize the individual variances of thickness. By focusing on these ROIs, we were able to reliably examine whether individual variances of FA, independent of thickness, in the language-related pathways reflect the individual linguistic abilities. In the present study, we focused on individual differences in L2 acquisition, because varieties in L2 abilities are usually larger than those in L1.

Here we recruited Japanese students who had learned English as L2, and assessed their individual syntactic abilities using a syntactic error-detection (Syn) task in English. The syntactic errors were basically related to the argument structures of English verbs. In some linguistic theories, it is assumed that "lexical entries contain at least some syntactic information, in addition to the phonological and semantic information that surely must be present" (Chomsky, 1995). Previous functional MRI studies have employed tasks for argument structures, revealing the selective activations for syntactic decisions in the left inferior frontal gyrus (IFG) (Suzuki and Sakai, 2003; Raettig et al., 2010). Magnetoencephalography (MEG) studies have also shown selective responses in the left IFG to verbs depending on argument structures, thereby controlling the effects of lexico-semantic factors (Iijima et al., 2009; Inubushi et al., 2012; Iijima and Sakai, 2014). Cortical regions other than the left IFG may also contribute to syntactic processing. It has been reported that damage to the left temporo-parietal cortex including the middle/superior temporal gyri and the angular gyrus is related to the decreased performances of sentence-picture matching tasks with several sentence types (e.g., active, passive, and cleft sentences) (Dronkers et al., 2004; Thothathiri et al., 2012; Magnusdottir et al., 2013). However, words used in these tasks may also have pragmatic or semantic effects, and the functions of these regions may be explained by these additional effects. Returning to the functions of the left IFG in syntactic processing,

we have clarified its critical involvement not only for L1, but for L2. In our previous functional MRI study (Sakai et al., 2009), we used the same tasks as in the present study, including a spelling error-detection (Spe) task to control semantic and related cognitive factors. We have demonstrated that the individual accuracy of the Syn task was positively correlated with activations of the dorsal triangular part (F3t) of the left IFG. Moreover, our voxel-based morphometry (VBM) study with the same tasks showed that the accuracy of the Syn task, but not of the Spe task, was correlated with the leftward lateralization of a single region's volume in the F3t (Nauchi and Sakai, 2009). These functional MRI and VBM studies have provided a consistent view that the individual differences both in the function and anatomy of the left IFG were related to the individual syntactic abilities. Regarding connectivity including regions adjacent to the left IFG, resting-state connectivity between the left anterior insula/frontal operculum and the left superior temporal region has been reported to show a positive correlation with other L2 abilities including lexical retrieval (Chai et al., 2016). In such resting-state connectivity, however, the *functional* roles of those pathways remain unknown. It would be important to examine the correlation between behavioral metrics and any properties of anatomically or functionally identified pathways. Given our previous studies suggesting the critical involvement of the left IFG to syntactic processing in L2, a next question is to clarify which of structural properties of white matter fibers connecting the left IFG reflect individual syntactic abilities.

Recent diffusion MRI studies have anatomically identified several dorsal and ventral pathways connecting inferior frontal regions with other cortical regions supporting language functions (Catani et al., 2012). Our recent study, using functional/diffusion MRI as well as functional connectivity analyses, showed that the left dorsal pathway including the Arcuate/SLF transmitted information from the left IFG to the supramarginal gyrus, where syntax-selective activations were observed (Ohta et al., 2013). On the other hand, a lesion study reported that disconnection in either the dorsal or ventral pathway was associated with

increased errors in a sentence-picture matching task, suggesting that both the dorsal and ventral pathways are critical for syntax (Griffiths et al., 2013). In order to reveal the roles of connectivity including the left IFG in L2 acquisition, it is thus important to separate syntactic abilities from other linguistic or general ones, and to further clarify whether the structural properties of the dorsal and/or the ventral pathways actually reflect individual syntactic abilities. In the present study, we focus on the Arcuate of the dorsal pathway, as well as the inferior fronto-occipital fasciculus (IFOF) of the ventral pathway, and hypothesize that structural properties of the *left* Arcuate reflect the individual syntactic abilities in L2. We recruited native Japanese speakers, who had studied English as L2 for about 5 years, and assessed their syntactic abilities in L2 based on the accuracy of the Syn task (Experiment I).

We further examined to what extent L2 abilities and the brain structures are associated with factors that are shared among individuals (e.g., a particular educational environment), and with those that are not shared among individuals (e.g., individual aptitudes for multiple languages). Because monozygotic twins share genetic and some environmental factors, twin studies can be used to explore the extent to which brain structure or function is influenced by shared genetic and environmental determinants. A monozygotic twin study has reported that the volume of each of four lobes was strongly affected by shared genetic/environmental factors, while the sulcal/gyral curvature and the surface-to-volume ratio were more prone to be affected by non-shared factors (White et al., 2002). Monozygotic twins have also been studied to reveal the influence of shared factors on sulcal patterning using a graph matching method (Im et al., 2011). A possible direction of research includes the following two steps: (i) identifying structural properties and behavioral performances that reflect specified functions (as in our Experiment I), and (ii) examining the correlations within monozygotic twin pairs for each of the structural and behavioral parameters to explore the extent of their dependence on shared factors. We assessed the behavioral correlations within monozygotic twin pairs (Experiment II), thereby examining the extent to which the accuracy

and reaction times (RTs) in the Syn/Spe tasks in L2, as well as a verbal fluency in L1, are associated with shared factors; we do not intend to dissociate genetic and environmental factors here. We also examined whether the white matter properties, i.e., the thickness and FA, were associated with shared factors. Our present study with monozygotic twins, together with appropriate methods to examine the structural basis of syntactic abilities, should shed new light on the neural framework underlying individual differences in L2 acquisition.

Here we conducted diffusion MRI acquisition and fiber tracking by using *q*-ball imaging. The *q*-ball imaging technique, a *q*-space analysis of diffusion MRI with a relatively high *b*-value (e.g., 4000 s/mm^2), was developed for improved reconstruction of the fiber tracts (Tuch, 2004). On the other hand, the *b*-value was often lowered to 1000 s/mm^2 for a better estimation of FA (Jones and Basser, 2004). For high *b*-values, the assumption that the diffusion signal decay is monoexponential may not be suitable (Clark and Le Bihan, 2000), but such a difference in FA is smaller with a 3T scanner than with a 1.5 T scanner (Chung et al., 2013). In the present study we used a 3T scanner and conducted a pilot experiment to examine the effects of imaging parameters.

2.2. Materials and methods

2.2.1. Experiment I

2.2.1.1. Participants

We recruited senior high-school students in their fifth academic year at the Secondary Education School attached to the Faculty of Education of the University of Tokyo, to which twins are preferentially admitted for educational research. We set the following basic inclusion criteria: (i) right-handedness as assessed by the Edinburgh inventory (Oldfield, 1971), (ii) no history of neurological or psychiatric diseases, and (iii) native Japanese speakers whose age of first exposure (AOE) to English in formal education was 12 or 13 years old (a condition met by the majority of students in this school). As regards the first

criterion, participants with a negative laterality quotient (LQ) or those with a potential history of change in handedness were dropped, resulting in a relatively strong right-handed population (LQ > 50 in most cases). In Experiment I, for each twin pair who met these criteria, the one exhibiting the higher score for the Spe task was entered into the analysis to avoid a double count of twins with potentially similar characteristics. There were 47 participants who met these conditions.

We set two additional criteria in Experiment I: (iv) longer RTs for Syn than for Spe; and (v) performance accuracy in the upper half of participants for Spe (80% or higher). The fourth criterion was necessary because shorter RTs for Syn indicated the possible abandonment of the more demanding task; two participants were dropped for this reason. Given that the participants in the present study did not receive any training sessions, it was necessary to set the fifth criterion in order to match the task accuracy of Syn [both mean and standard deviation (SD)] with that *after the training* reported in our previous studies (Nauchi and Sakai, 2009; Sakai et al., 2009); 19 participants were dropped for this reason. When the fifth criterion was not applied, the accuracy of Syn for 20 out of the 47 participants was lower than 55% (i.e., near the level of chance); such a population was not appropriate for examining individual syntactic abilities. Twenty-six participants met the five criteria, and demographic details of these participants are shown in Table 2.1: age, AOE to English, duration of exposure (DOE) to English, and LQ. Written informed consent was obtained from all participants as well as from their guardians. The study was approved by the Secondary Education School and by the Institutional review board of the University of Tokyo, Komaba.

Table 2.1. Demographic details of the participants in each experiment.

	N	Age	AOE	DOE	LQ
Experiment I	26	17 ± 0.3	12.6 ± 0.3	4.5 ± 0.2	81 ± 18
	(15 females)	(16−17)	(12−13)	(4.3−4.8)	(38−100)
Experiment II	24	17 ± 0.3	12.5 ± 0.3	4.5 ± 0.2	87 ± 13
	(18 females)	(16−17)	(12−13)	(4.3−4.8)	(63−100)

The number of participants (N), age, age of first exposure (AOE) to English, duration of exposure (DOE) to English, and laterality quotient (LQ) are shown as the mean ± SD. Ranges are shown in brackets.

2.2.1.2. Stimuli and tasks

We visually presented stimuli, which were English sentences identical to those used after training sessions in our previous studies (Nauchi and Sakai, 2009; Sakai et al., 2009). In brief, we selected 42 high-frequency English verbs (20 transitive and 22 intransitive, including 12 unergatives and 10 unaccusatives (Yusa, 2003)), and used them to make sentence stimuli with other high-frequency words; each verb was used 1–3 times to test different properties of these verbs. Each stimulus consisted of a key sentence and its associated sentence, the latter of which was either syntactically normal or anomalous in the Syn task. It is generally hard for L2 learners to acquire correct argument structures and their associated syntactic knowledge. For instance, Japanese students learning English as L2 tend to omit an object and to accept such incorrect sentences (*) as *"Do you often meet Mary?—*Yes, I often meet*," since objects of transitive verbs can be omitted quite freely in many languages including Japanese. In this way, our task properly tested the structural properties of verbs. The stimuli contained several other types of errors, such as an incorrect use of intransitive verbs ("*You can kill many monsters in the game.—*You will die them in the game.*") or subject drop ("*John often comes to my house.—*Soon will come.*"), requiring the relational properties of the constituents.

Because these ungrammatical sentences cannot be judged as incorrect by semantic information alone, the accuracy of the Syn task appropriately reflected individual syntactic abilities. In the Spe task, a typographical error was included to test the English orthography of words. Moreover, the sets of sentences were basically the same in the Syn and Spe tasks, and multiple words were naturally related to one another to understand the meanings of these sentences in both tasks.

All behavioral data were acquired outside the MR scanner. At the initiation of every trial of 7 s, a cue, indicating whether the task was Syn or Spe, was shown for 400 ms. Next, a set of sentence stimuli was shown for 6400 ms. The participants were instructed to read the sentences silently, and to indicate whether or not the sentences contained an error by pushing one of two buttons. For each task, *five trials* were consecutively performed as a task block for better concentration on each type of task, and these task blocks were alternately performed. The initial order of task blocks was counterbalanced across participants, thereby eliminating any effects of the order of presentation [i.e., Syn—Spe—Syn… or Spe—Syn—Spe…]. Participants performed 50 trials, and after a short break, they performed another 50 trials, accomplishing 100 trials in total (50 each for Syn and Spe; 50 each for correct and incorrect stimuli, completely randomized). As in the previous studies with the same tasks, trials with no response within the allowed time were included as incorrect ones in calculating the accuracy of each task; a response within 1.5 s was regarded as the response to the trial just one before. As regards RTs, only trials with correct responses were included. After the task instruction and brief explanation, the participants received 10 training trials (five for Syn, five for Spe). Additional 10 trials were provided for some participants who needed to adapt to the time constraint of the tasks.

To test general skills in word processing, a verbal fluency task in L1 (Japanese) was performed. For a given single letter in hiragana, participants were asked to write as many words (except proper nouns) starting from this letter as possible within 60 s. The verbal

fluency was separately tested in three trials with different letters, which consisted of an easier letter (*sa*) and two harder letters (*ri* and *nu*).

2.2.1.3. MR image acquisition

The scans were conducted on a 3.0 T scanner (Signa HDxt; GE Healthcare, Milwaukee, WI, USA) with an 8-channel phased array head coil. We set the axial sections of the brain in parallel with its anterior to posterior commissure (AC-PC) line. We scanned 50 axial slices that were 3-mm thick without a gap, covering the whole brain, with a diffusion-weighted spin-echo echo-planar imaging sequence (*b*-value = 4000 s/mm^2, repetition time (TR) = 17 s, echo time (TE) = 110 ms, field of view (FOV) = 192 × 192 mm^2, resolution = 3 × 3 mm^2, number of excitations = 1). For this scanning, a single image without diffusion-weighting (b0) was initially acquired, and then diffusion-weighting was isotropically distributed along 60 diffusion-encoding gradient directions. After completion of the diffusion MRI session, high-resolution T1-weighted images of the whole brain (192 axial slices, 1 × 1 × 1 mm^3) were acquired from all participants with a fast spoiled gradient recalled acquisition in the steady state sequence (TR = 10 ms, TE = 4 ms, FA = 25°, FOV = 256 × 256 mm^2).

2.2.1.4. Diffusion MRI data analyses

We used *q*-ball imaging (Tuch, 2004), a *q*-space analysis of diffusion MRI, and performed data analyses using FSL [Oxford Centre for Functional MRI of the Brain's (FMRIB) Software Library 4.1.9; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/] and FDT 2.0 (FMRIB's Diffusion Toolbox) (Smith et al., 2004). Diffusion-weighted images were first resliced to an isotropic voxel size of 1 mm^3, and then eddy current distortions and motion artifacts were corrected using affine registration to the b0 image. We then extracted the brain shape from the b0 image, and created a binary mask image (zero for the outside of the brain) for each participant to calculate the diffusion tensor, three eigenvectors, and FA in each voxel inside

the brain. Fiber orientations were estimated for the constant solid angle orientation distribution functions (ODFs) using the spherical harmonic decomposition and residual bootstrapping algorithm included in the QBOOT (*q*-ball ODFs and residual bootstrap) toolbox in FSL (Sotiropoulos et al., 2013).

We performed tractography by a single-seed approach in the native space using the probtrackx tool of FDT; we basically followed the procedures described previously (Ohta et al., 2013). Probabilistic fiber tracking was initiated from each voxel within a seed mask to generate 5000 streamline samples in any direction, with a step length of 0.5 mm, a maximum number of steps of 2000, a curvature threshold of 0.2 (±78.5°), and a loopcheck option. Each voxel value represented the total number of streamline samples passing through that voxel, making a connectivity distribution. The connectivity probability maps were then created by dividing the connectivity distribution with a waytotal value, i.e., the total number of generated streamline samples from the seed mask that had reached two waypoint masks. To remove any spurious connections, tracts were thresholded to include only voxels that had connectivity probability values of at least 0.2%. Thresholded tracts were then binarized using the fslmaths function of FSL. In order to prevent the misplacement of seed and waypoint masks, all tracts were inspected to confirm that they were not touching the peripheral edges of the masks.

2.2.1.5. Identification of the Arcuate and IFOF

By using FMRIB's Linear Image Registration Tool (FLIRT) on FSL, the b0 image was first coregistered to the individual T1-weighted image for each participant, and the T1-weighted image was spatially normalized to the Montreal Neurological Institute (MNI) space by using both affine and nonlinear transformation with FLIRT and FMRIB's Nonlinear Image Registration Tool (FNIRT). To reliably set the *center* of a seed mask, we placed a "bottleneck mask" (a sphere of 10-mm radius) at a region (±35, −35, 30) in the MNI space where a

narrower portion of the Arcuate was expected to pass, with reference to the atlases of white matter pathways (Catani and Thiebaut de Schotten, 2008; Oishi et al., 2011). With the transformation matrices and estimated deformation fields, the bottleneck mask was transformed back to the native space of each participant (orange ellipses in Figure 2.1A). On a participant's color-coded map of fiber bundles, we confirmed that the center-of-gravity of the transformed bottleneck mask was included in the tract's cross-section; the transformed point was manually adjusted to the center of the tract within a few millimeters [mean ± SD; Experiment I: 3.8 ± 1.3 mm (left), 4.4 ± 2.1 mm (right); Experiment II: 3.8 ± 1.6 mm (left), 4.0 ± 1.1 mm (right)]. The amount of adjustments had no hemispheric difference according to the paired *t*-tests in Experiment I ($t(25) = 1.3$, $P > 0.1$) or in Experiment II ($t(23) = 0.4$, $P > 0.6$). On a coronal slice, we generated a *seed* mask of 25-mm square (the red square in Figure 2.1B), whose center was placed at the transformed point. The resultant seed mask included the entire cross-section of the Arcuate. These semi-automatic procedures allowed us to define seed masks in a consistent manner among participants.

To reliably extract the Arcuate alone, we used two *waypoint* masks of 30-mm squares (blue squares in Figure 2.1B) that were large enough to include the tract's cross-sections. A coronal mask was placed 20-mm anterior to the seed mask in the native space, and an axial mask was placed 20-mm inferior to the seed mask's center. For the coronal mask, its lowest edge was parallel to that of the seed mask at the same *z*-coordinates; both masks' centers also had the same *x*-coordinates. The axial mask's center had the same *x*-coordinates as the seed mask, and the *y*-coordinate of the axial mask's center was initially set 10 mm posterior to the seed mask. By considering individual variability near the arching part of the pathways, we adjusted the *y*-coordinate within a few millimeters [Experiment I: 2.1 ± 2.7 mm (left), –1.3 ± 3.7 mm (right); Experiment II: 2.8 ± 2.5 mm (left), –1.2 ± 2.7 mm (right)]. The amount of adjustments had hemispheric differences in Experiment I ($t(25) = 5.1$, $P < 0.05$) and in Experiment II ($t(23) = 7.3$, $P < 0.05$), indicating that the curvature of the

Arcuate was different in each hemisphere. We also used a set of exclusion masks in three planes to avoid the potential inclusion of fibers other than the Arcuate: a coronal mask (15-mm square) to exclude the IFOF, an axial mask (30-mm square) to exclude the corona-radiata, and a mid-sagittal plane to exclude commissural fibers. The coronal exclusion mask was the same as the seed mask for the IFOF (see below), and the axial exclusion mask was placed 10 mm superior to the center of the coronal exclusion mask, and individually adjusted in the axial plane to cover a narrower portion of the central vertical fibers on the color-coded map of fiber bundles.

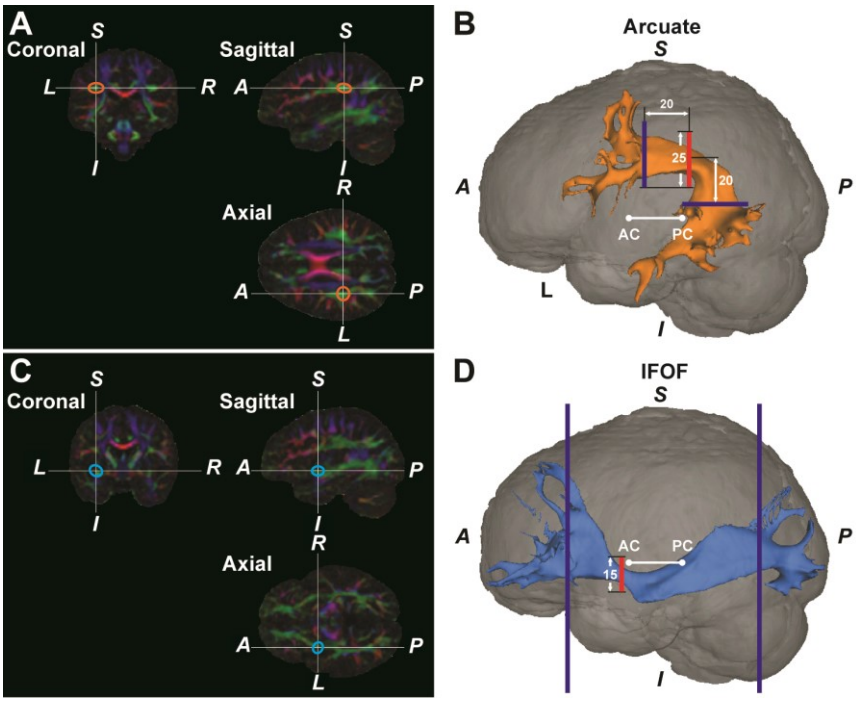

Figure 2.1. The Arcuate and inferior fronto-occipital fasciculus (IFOF) reconstructed by diffusion magnetic resonance imaging (MRI) in a typical participant. (A) A color-coded map of fiber bundles in the native space of a participant. In the Montreal Neurological Institute (MNI) space, we placed a "bottleneck mask" (a sphere of 10-mm radius) at a region where a narrower portion of the pathway was expected to pass. The mask was then transformed back to the native space. The orange ellipses denote the approximate outline of the transformed bottleneck mask on the left Arcuate. (B) The reconstructed left Arcuate. On a coronal slice, we generated a seed mask (shown in red), whose center was placed at the center of the transformed bottleneck mask on the left Arcuate. The waypoint masks are shown in blue. (C) A color-coded map of fiber bundles for the same participant. The blue ellipses denote the outline of the transformed bottleneck mask on the left IFOF. (D) The reconstructed left IFOF. On a coronal slice, we generated a seed mask (shown in red), whose center was placed at the center of the transformed bottleneck mask on the left IFOF. *A*, anterior; *P*, posterior; *S*, superior; *I*, inferior; *L*, left; *R*, right; AC, anterior commissure; PC, posterior commissure.

As regards the IFOF, we again placed a bottleneck mask (a sphere of 10-mm radius) at a region (±35, 0, −5) in the MNI space where a narrower portion of the IFOF was expected to pass, with reference to the atlases of white matter pathways. The bottleneck mask was transformed back to the native space of each participant (blue ellipses in Figure 2.1C). We confirmed that the center-of-gravity of the transformed bottleneck mask was included in the tract's cross-section; the transformed point was manually adjusted to the center of the tract within a few millimeters [Experiment I: 4.9 ± 1.3 mm (left), 5.1 ± 2.0 mm (right); Experiment II: 4.5 ± 1.1 mm (left), 4.7 ± 2.0 mm (right)]. The amount of adjustments had no hemispheric difference according to the paired *t*-tests in Experiment I ($t(25) = 0.5$, $P > 0.5$) or in Experiment II ($t(23) = 0.3$, $P > 0.6$). On a coronal slice, we generated a seed mask of 15-mm square (the red square in Figure 2.1D), whose center was placed at the transformed point. To include longer tracts connecting both the IFG and occipital regions, we used waypoint masks at two coronal planes in each individual hemisphere: one at the head of the caudate nucleus, and the other at the parieto-occipital sulcus (blue squares in Figure 2.1D) (Forkel et al., 2014). We also used a set of exclusion masks in two planes to avoid the potential inclusion of fibers other than the IFOF: a coronal mask to exclude the Arcuate and superior fronto-occipital fasciculus (Forkel et al., 2014), and a mid-sagittal mask to exclude commissural fibers. The coronal exclusion mask covered only the upper half of the native MR image, and was placed at the anteroposterior midpoint of the image.

2.2.1.6. ROI selection to minimize individual variances in the fiber tract's thickness

The *thickness* of the fiber tract was defined as the number of voxels (voxel size, 1 mm^2) at a certain section of the tracked fibers. Here we chose coronal sections, because the portions without branching in the Arcuate or IFOF were nearly horizontal. We measured the thickness along the Arcuate in each hemisphere for the length of 35 mm, i.e., cross-sections L_1~L_{35} of the upper half of the image, corresponding to the 94th~128th coronal slices from the anterior

end of the image (see Figure 2.2A). Within a tract segment of a fixed length (e.g., 20 mm), we calculated the SD of the tract's thickness, and slid the segment between L_1 and L_{35}. Among these segments, the segment with a minimal SD, i.e., with the most uniform thickness, was selected as a ROI. To objectively determine its appropriate length or *ROI size*, we varied it from 10 to 30 mm, and the "mean thickness" was calculated among the selected cross-sections. The mean thickness was further averaged among the participants (Figure 2.2B, C). The averaged thickness started to increase when the ROI sizes exceeded 20 mm in both hemispheres (Figure 2.2B), indicating the contribution from branching portions. In the left hemisphere, the SD of the mean thickness among participants was lowest for the ROI size of around 20 mm (Figure 2.2C). We thus chose a ROI size in the range of 19–21 mm, in proportion to the individual size of the brain. For 47 out of 52 tracts in all hemispheres tested, the final ROIs did not reach the end of L_1 or L_{35}; thus the search volume of L_1~L_{35} was large enough to determine most ROIs.

As regards the IFOF, we measured thickness for the length of 70 mm, i.e., cross-sections L_1~L_{70}, corresponding to the 84th~153th coronal slices from the anterior end of the image (see Figure 2.2D). Within a tract segment of a fixed length (e.g., 15 mm), we calculated the SD of the tract's thickness, and slid the segment between L_1 and L_{70}. Among these segments, the segment with a minimal SD was selected as a ROI. To objectively determine an appropriate length or ROI size, we varied it from 10 to 30 mm, and calculated the "mean thickness." The averaged thickness among the participants started to increase when the ROI sizes exceeded 15 mm in the left hemisphere (Figure 2.2E). In both hemispheres, the SD of the mean thickness among participants was lowest for the ROI size of around 15 mm (Figure 2.2F). We thus chose a ROI size in the range of 14–16 mm, in proportion to the individual size of the brain. For all tracts tested, the final ROIs did not reach the end of L_1 or L_{70}.

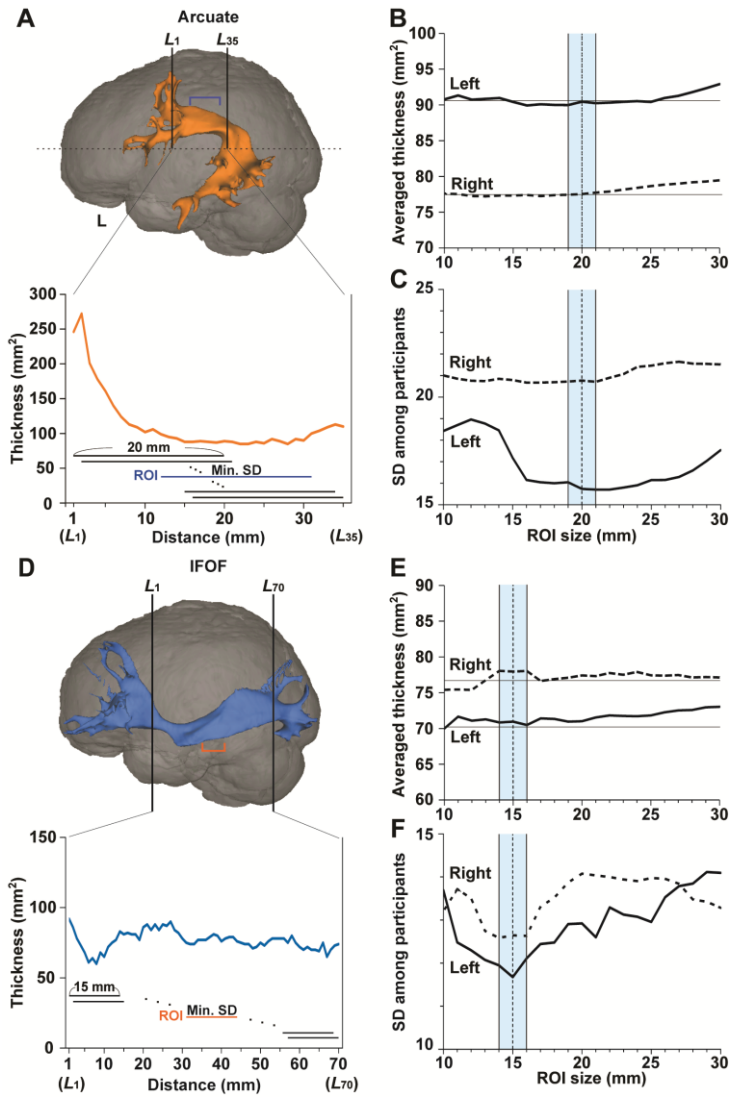

Figure 2.2. Region of interest (ROI) selection to minimize individual variances in the fiber tract's thickness. (A) A ROI (a segment shown by blue lines) in the left Arcuate reconstructed for the participant shown in Figure 2.1. Thickness was defined as the number of voxels (voxel size, 1 mm^2) at each coronal section of the tracked fibers. In the lower graph, the tract's thickness is shown at each of the L_1~L_{35} cross-sections for the upper half of the brain (above the horizontal dashed line). Within a tract segment of a fixed length (e.g., 20 mm), we calculated the standard deviation (SD) of the tract's thickness, and slid the segment between L_1 and L_{35}. Among these segments, the segment with a minimal SD, i.e., with the most uniform thickness, was selected as a ROI. (B, C) The averaged thickness and SD for various ROI sizes in the Arcuate. The mean thickness in the ROI was further averaged, and the SD was obtained among the participants. The averaged thickness or SD in the left and right hemispheres is shown by the solid and dashed lines, respectively. To minimize the SD, we chose a ROI size in the range of 19–21 mm (blue-shaded), in proportion to the individual size of the brain. (D) A ROI (a segment shown by orange lines) in the left IFOF reconstructed for the same participant. In the lower graph, the tract's thickness is shown at each of the L_1~L_{70} cross-sections. Among these segments, the segment with a minimal SD was selected as a ROI. (E, F) The averaged thickness and SD for various ROI sizes in the IFOF. The averaged thickness or SD in the left and right hemispheres is shown by the solid and dashed lines, respectively. We chose a ROI size in the range of 14-16 mm (blue-shaded), in proportion to the individual size of the brain.

2.2.1.7. FA analyses

As described in the above sections, ROIs were determined on the basis of the connectivity probability and resultant thickness. After selecting ROIs, voxels with FA values of 0.2 or higher were selected, and this thresholding was used only in FA analyses. The thresholding has been widely used in previous studies to avoid the uncertainty of anisotropy measurements in the peripheral regions of fibers. We calculated the mean FA within the ROI, and examined the relationships between the accuracy of the Syn task and FA in each pathway after removing the effects of the accuracy of Spe, gender, and LQ, using the ppcor (partial and semi-partial correlation) package (https://cran.r-project.org/web/packages/ppcor/) in R software (http://www.r-project.org/).

2.2.2. Experiment II

In Experiment II, only monozygotic twin pairs (24 twins) were included (Table 2.1); we set the first three inclusion criteria (i)–(iii) used in Experiment I. Note that the criterion for the accuracy of Spe was *not* employed in Experiment II. Eight participants were analyzed in both Experiment I and II, but these analyses were independent and did not constitute selection and selective evaluation analyses. As in Experiment I, the participants performed the Syn, Spe, and verbal fluency tasks, and MR imaging acquisition was conducted with the same parameters. The Arcuate and IFOF were identified in each individual hemisphere, as described above. We examined correlations within twin pairs separately for the behavioral data and the structural (FA/thickness) data. To examine the difference between correlations, we statistically compared correlation coefficients, using the cocor (comparing correlations) package (https://cran.r-project.org/web/packages/cocor/) in R software (Diedenhofen and Musch, 2015).

2.2.3. Effects of different imaging parameters on FA

In a pilot experiment for 13 participants (including four pairs of monozygotic twins) from Experiment I/II, diffusion MRI data were acquired with a *b*-value of 1000 s/mm^2 (TR = 17 s, TE = 90 ms, FOV = 256 × 256 mm^2, resolution = 2 × 2 mm^2, number of excitations = 1), in addition to diffusion MR data with a *b*-value of 4000 s/mm^2 (see "MR image acquisition" section). We scanned 67 axial slices that were 2-mm thick without a gap, covering the whole brain. A single b0 image was initially acquired, and then diffusion-weighting was isotropically distributed along 60 diffusion-encoding gradient directions, as in Experiment I/II. In our MRI system, we were able to set a smaller voxel size for a lower *b*-value of 1000 s/mm^2 (2 mm^3) than that for a *b*-value of 4000 s/mm^2 (3 mm^3). We confirmed that FA from the imaging with a *b*-value of 4000 s/mm^2, which has been used for improved reconstruction of the fiber tracts (see "Introduction" section), was consistent with FA obtained with a lower *b*-value and higher resolution. By using FDT 3.0 and FLIRT on FSL 5.0.4, the individual b0 image with the *b*-value of 1000 s/mm^2 was coregistered to that with the *b*-value of 4000 s/mm^2. With the same transformation matrices, an individual FA map with a *b*-value of 1000 s/mm^2 was also spatially transformed for each participant. We then compared the FA values in the left Arcuate obtained with different *b*-values, within the tract reconstructed by *q*-ball imaging with the *b*-value of 4000 s/mm^2.

2.3. Results

2.3.1. Experiment I

2.3.1.1. Behavioral data

Behavioral data of the Syn and Spe tasks are shown in Table 2.2. For the verbal fluency task, we evaluated each individual's performance by adding together the number of words produced for all trials (mean ± SD: 26.2 ± 6.3). The performance of the verbal fluency task

was not correlated with any of the accuracy/RTs in Syn or Spe ($P > 0.2$). The accuracy of Syn was significantly lower than that of Spe ($t(25) = 11.4$, $P < 0.0001$), and the RTs of Syn were significantly longer than those of Spe ($t(25) = 12.6$, $P < 0.0001$), indicating that the Syn task was more demanding than the Spe task.

Table 2.2. Behavioral data of the participants.

	Accuracy (%)		RTs (ms)	
	Syn	Spe	Syn	Spe
Experiment I	61 ± 11	86 ± 4.9	4778 ± 383	4324 ± 410
	(42−88)	(80−100)	(3989−5378)	(3503−5006)
Experiment II	60 ± 8	82 ± 9.2	4806 ± 489	4316 ± 529
	(44−73)	(68−100)	(3989−5877)	(3503−5587)

Only trials with correct responses were included for the reaction times (RTs). Data are shown as the mean ± SD. Ranges are shown in brackets.

If a behavioral parameter reflected general cognitive factors, such as task difficulty and reading proficiency, which were commonly involved in the Syn and Spe tasks with some differences in degree, then that parameter should show a correlation between Syn and Spe among the participants, when considered their individual differences regarding these factors. For example, a slower reader would show longer RTs for both Syn and Spe. Therefore, as a contraposition, if another behavioral parameter did not show such a correlation between Syn and Spe among the participants, that parameter reflects at least distinct factors required in each task. We thus examined the relationships between the performances of the two tasks. The accuracy of Syn and that of Spe did not show a significant correlation ($r = 0.20$, $P = 0.32$; Figure 2.3A), suggesting that the accuracy mainly reflected distinct abilities required in each task. On the other hand, the RTs of Syn and Spe showed a highly positive correlation ($r = 0.89$, $P < 0.0001$; Figure 2.3B), indicating that the RTs reflected general cognitive factors common to both tasks. We observed that the correlation coefficient between RTs was

significantly larger than that between the accuracy ($Z = 3.1$, $P < 0.05$) by using the cocor (comparing correlations) package in R software, confirming the different natures of the accuracy and RTs. These results indicate that the *accuracy* of each task for the participants reflected individual abilities required in each task.

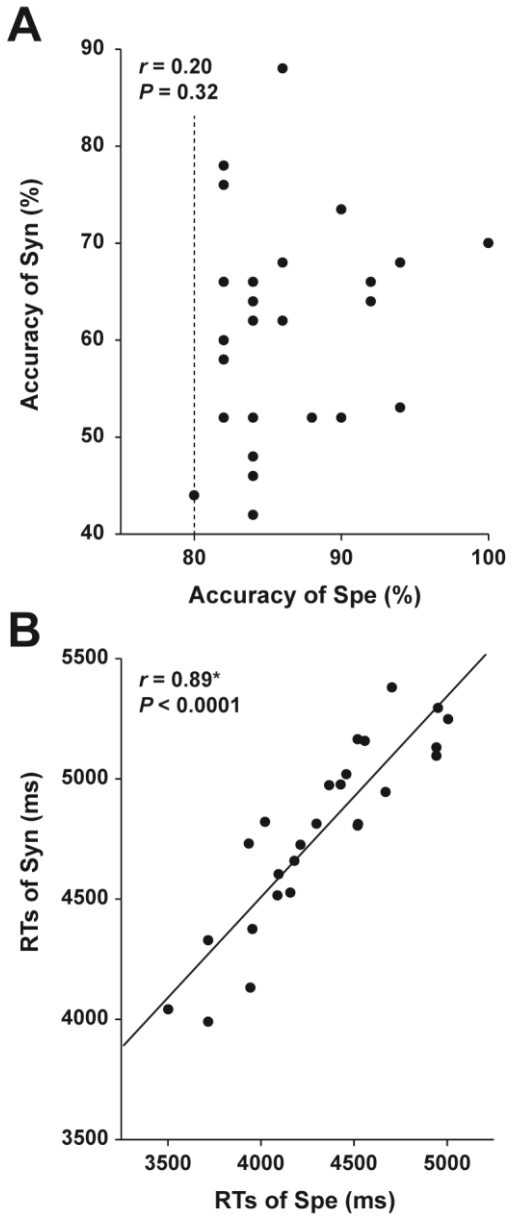

Figure 2.3. Independence of the accuracy of the Syn and Spe tasks. Individual behavioral parameters in Experiment I are plotted for comparing the two tasks. The dotted line indicates an inclusion criterion of Spe employed in Experiment I. The accuracy of Syn and Spe showed no significant correlation (A), while reaction times (RTs) of Syn and Spe were highly correlated (B). An asterisk at *r* indicates corrected $P < 0.05$.

2.3.1.2. Structural properties of the Arcuate and IFOF

Structural properties obtained in each ROI of the Arcuate and IFOF are shown in Table 2.3. Within the tested group, the distribution was not significantly different from the normal distribution for the mean FA and thickness in each ROI of the four pathways [the Shapiro-Wilk test (Shapiro and Wilk, 1965); $P > 0.08$]. As regards the mean FA, a two-way repeated-measures analysis of variance (rANOVA) [pathway (Arcuate, IFOF) × hemisphere (left, right)] showed a significant main effect of pathway ($F(1,25) = 54.2$, $P < 0.001$), but with neither a main effect of hemisphere ($F(1,25) = 0.86$, $P = 0.36$) nor an interaction ($F(1,25) = 0.73$, $P = 0.4$). According to a paired *t*-test (two-tailed) in each hemisphere, the mean FA in the ROI of the IFOF was significantly higher than that of the Arcuate (left: $t(25) = 5.3$, $P < 0.001$; right: $t(25) = 5.0$, $P < 0.001$). With respect to the mean thickness, an rANOVA showed a significant main effect of pathway ($F(1,25) = 9.0$, $P < 0.01$) and an interaction ($F(1,25) = 26.9$, $P < 0.001$), but no main effect of hemisphere ($F(1,25) = 1.2$, $P = 0.3$). Paired *t*-tests showed that the mean thickness in the ROI of the left Arcuate was significantly larger than that of the right Arcuate ($t(25) = 3.7$, $P < 0.005$), while that of the right IFOF was marginally larger than that of the left IFOF ($t(25) = 2.1$, $P = 0.049$; significance level at $\alpha = 0.025$, Bonferroni-corrected). These results indicated that the Arcuate and IFOF were different in terms of both macroscopic and microscopic properties, i.e., thickness and FA, respectively, even when the same objective methods for ROI selection were employed.

Table 2.3. Structural properties of the Arcuate and IFOF.

	Mean FA		Mean thickness (mm^2)		
	Left	Right	Left	Right	Difference
Arcuate	0.39 ± 0.02	0.39 ± 0.02	90 ± 16	78 ± 21	$P < 0.005$
IFOF	0.44 ± 0.04	0.43 ± 0.04	71 ± 11	78 ± 13	$P = 0.049$

For each pathway of either hemisphere, the mean FA or mean thickness in each ROI was obtained in Experiment I. Data are shown as the mean ± SD among the participants ($N = 26$). Hemispheric difference was assessed by paired *t*-tests.

2.3.1.3. FA of the left Arcuate as an indicator of individual performances for the Syn task

Within the tested group, the distribution was not significantly different from the normal distribution for the behavioral performances (the accuracy of Syn, and the number of produced words in the verbal fluency task), as well as for the mean FA in each ROI of the four pathways (the Shapiro-Wilk test; $P > 0.1$). For each of the four pathways, we performed partial correlation analyses between the standardized mean FA and the standardized accuracy of Syn (two-tailed, significance level at $\alpha = 0.0125$, Bonferroni-corrected), removing the effects of other major independent factors: the standardized accuracy of Spe, gender, and LQ (there were no correlations among these factors: $r < 0.3$, $P > 0.3$). The mean FA of the left Arcuate showed a significant positive correlation with the accuracy of Syn ($r = 0.54$, $P = 0.007$; Figure 2.4A). In contrast, the mean FA in the ROI of the right Arcuate was not significantly correlated with the accuracy of Syn ($r = 0.06$, $P = 0.8$; Figure 2.4B). As regards the IFOF, the mean FA in either hemisphere was not significantly correlated with the accuracy of Syn (left: $r = -0.34$, $P = 0.1$; right: $r = -0.09$, $P = 0.7$; Figure 2.4C, D). Additionally, considering that FA (≥ 0.2) in each ROI might not be normally distributed, we tested all analyses above for the median FA instead of the mean, and replicated the overall results (a partial correlation between the median FA in the left Arcuate and the accuracy of Syn; $r = 0.56$, $P = 0.003$). Regarding the accuracy of Spe, we performed partial correlation analyses between the standardized mean FA and the standardized accuracy of Spe, thereby removing the effects of the standardized accuracy of Syn, gender, and LQ. The mean FA of the left Arcuate did not show a significant correlation with the accuracy of Spe ($r = 0.04$, $P = 0.9$). None of the other pathways showed a significant correlation with the accuracy of Spe ($P > 0.19$). Note that the range of the accuracy of Spe was limited due to the fifth inclusion criterion (see "Participants" section), and that the statistical power was lower. To further confirm the absence of the effects of the load associated with word processing, we also performed partial correlation analyses between the standardized mean FA and the

standardized performance of the verbal fluency task. The mean FA of the left Arcuate was not correlated with the performance of the verbal fluency task (r = 0.13, P = 0.5), and the other pathways did not show significant correlations (P > 0.5). These results demonstrate that FA of the left Arcuate can be a selective indicator of individual performances for the Syn task.

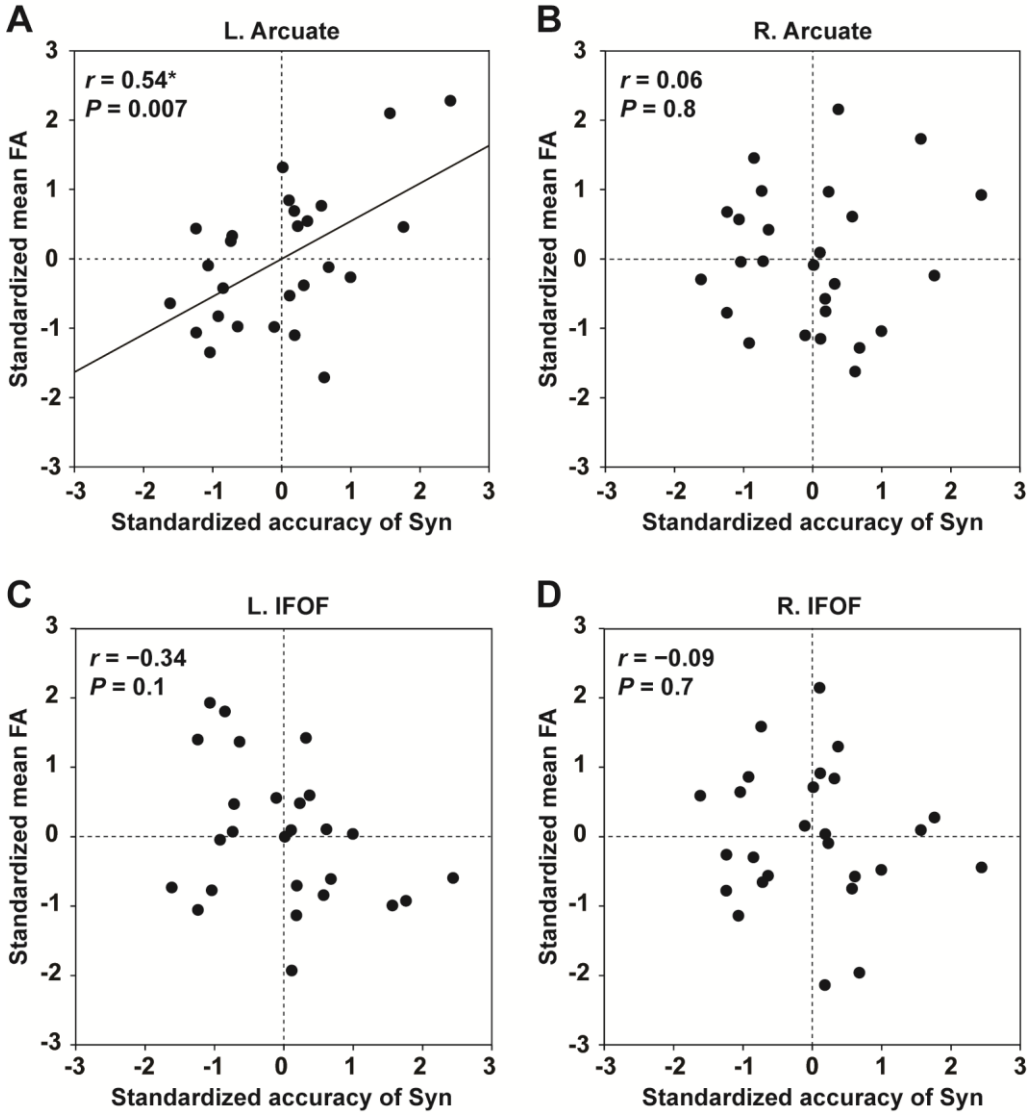

Figure 2.4. FA of the left Arcuate as an indicator of individual performances for the Syn task. The scatter plots of the standardized accuracy of Syn and the standardized mean FA in each ROI are shown for the left Arcuate (A), right Arcuate (B), left IFOF (C), and right IFOF (D). The effects of the standardized accuracy of Spe, gender, and laterality quotient (LQ) were removed. Uncorrected P values are also shown for each panel. Only the left Arcuate showed a significant correlation between FA and the accuracy of Syn.

2.3.2. Experiment II

2.3.2.1. Behavioral correlations within twin pairs

Behavioral data for monozygotic twins are shown in Table 2.2. Each monozygotic twin was randomly assigned as either twin A or twin B in a pair. Within the twin A or twin B group, the distribution of each of the task performances in L2 was not significantly different from the normal distribution (the Shapiro-Wilk test; $P > 0.1$). We examined correlations within twin pairs for the behavioral data (two-tailed, significance level at $\alpha = 0.0125$, Bonferroni-corrected). As regards the Syn task, the correlation of the accuracy within pairs was marginally correlated ($r = 0.66$, $P = 0.02$; Figure 2.5A), while the RTs were significantly correlated ($r = 0.78$, $P = 0.003$; Figure 2.5B). As regards the Spe task, the twins exhibited highly correlated performances in terms of accuracy ($r = 0.85$, $P = 0.0004$; Figure 2.5C) and RTs ($r = 0.96$, $P < 0.0001$; Figure 2.5D). Additionally, the correlation coefficient for the RTs of Spe was significantly larger than that for the accuracy of Syn ($P < 0.05$). With respect to the verbal fluency task, which was also normally distributed ($P > 0.09$), the total number of produced words was significantly correlated for the twins ($r = 0.81$, $P = 0.001$).

To further confirm that these correlations did not occur by chance, each twin was randomly assigned to a new pair to produce 10,000 permutations (Sakai et al., 2004). These random pairings resulted in a normal distribution (RTs of Syn: mean $r = -0.07$, SD = 0.27; accuracy of Spe: mean $r = -0.08$, SD = 0.29; RTs of Spe: mean $r = -0.08$, SD = 0.26; verbal fluency: mean $r = -0.08$, SD = 0.29), establishing that the above results were actually due to factors that each twin pair had in common.

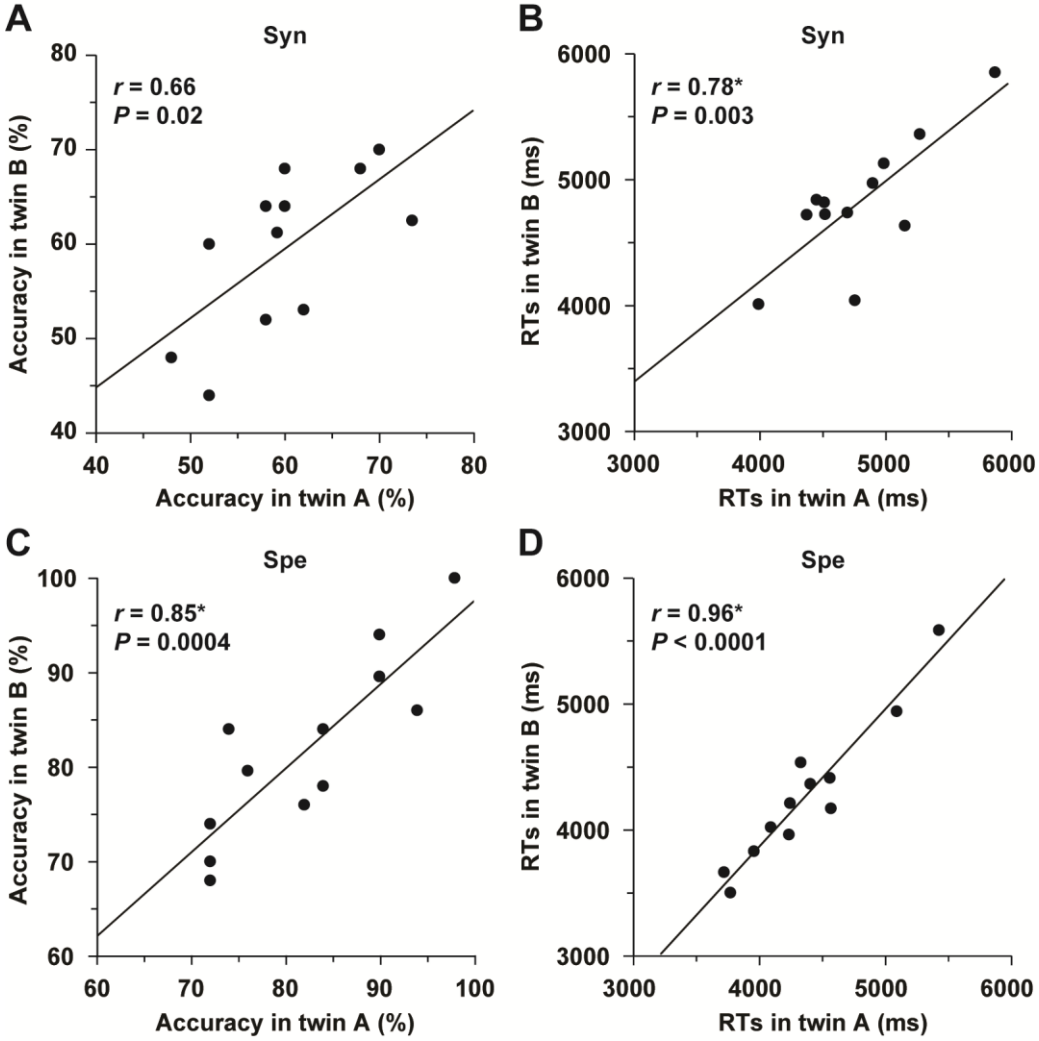

Figure 2.5. Behavioral correlations within twin pairs. (A–D) Correlations for accuracy/RTs of the Syn and Spe tasks are shown separately, and plotted for each pair of monozygotic twins. Each twin was randomly assigned as either twin A or twin B in a pair. The twin pairs showed significant correlations for task performances.

2.3.2.2. Structural correlations within twin pairs

Within the twin A or twin B group, the distribution of thickness/FA in the ROIs of the Arcuate and IFOF was not significantly different from the normal distribution (the Shapiro-Wilk test; $P \geq 0.05$). We analyzed the structural correlations within twin pairs for the Arcuate (two-tailed, significance level at $\alpha = 0.0125$, Bonferroni-corrected). As regards the mean thickness, the twins were highly correlated in the left hemisphere ($r = 0.79$, $P = 0.002$; Figure 2.6A), but not in the right hemisphere ($r = 0.57$, $P = 0.05$; Figure 2.6B). Random pairings for the

thickness of the left Arcuate resulted in a normal distribution (mean $r = -0.05$, SD = 0.29), confirming that the significant correlation was due to factors that each twin pair had in common. With respect to the mean FA, the correlation was not significant in either hemisphere (left: $r = 0.33$, $P = 0.30$; right: $r = 0.55$, $P = 0.06$; Figure 2.6C, D). For the IFOF, the correlation was not significant in either hemisphere as regards the mean thickness (left: $r = 0.19$, $P = 0.56$; right: $r = 0.05$, $P = 0.89$) or the mean FA (left: $r = 0.57$, $P = 0.05$; right: $r =$

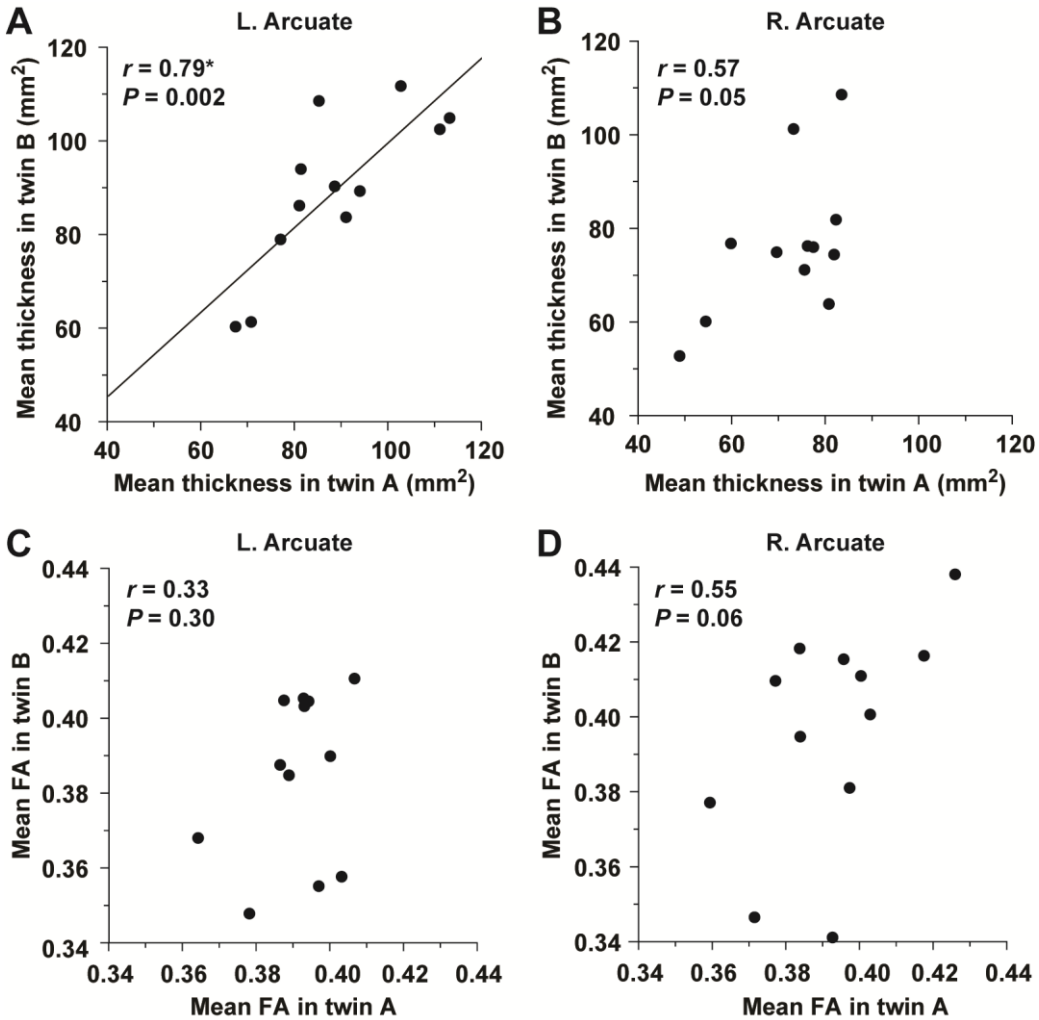

Figure 2.6. The structural correlation of the left Arcuate within twin pairs. The mean thickness in the ROI of the left (A) or right (B) Arcuate is plotted for each pair of monozygotic twins. The plot of the mean FA in the ROI of the left (C) or right (D) Arcuate is also shown. Note the significant correlation for the mean thickness of the left Arcuate alone.

−0.21, $P = 0.50$). Indeed, the correlation coefficient for the thickness of the left Arcuate was significantly larger than that of the left IFOF ($P < 0.05$).

In Experiment II, where the criterion for the accuracy of Spe was not employed, the accuracy of Syn and that of Spe were correlated ($r = 0.56$, $P < 0.01$). Because the assumption of independence between the accuracy of the two tasks in Experiment I did not apply in Experiment II, we considered that further correlation analyses among multiple metrics/ properties of behavior and structure (e.g., cross-correlation analyses between the accuracy of Syn in twin A and structural properties in twin B, and vice versa) were inconclusive in Experiment II.

2.3.3. Effects of the imaging parameters on FA

In the pilot experiment, we examined whether the FA values obtained with different *b*-values were correlated. As shown in Figure 2.7A, the overall profiles of FA averaged among the participants were similar for both *b*-values throughout the range of L_1~L_{35} in the left Arcuate, although FA itself was higher for 1000 s/mm^2. To examine the global tendency, we calculated the mean FA among the voxels in the entire range of L_1~L_{35} for each participant, and examined the correlation between the mean FA values at different *b*-values. As shown in Figure 2.7B, the mean FA with the higher *b*-value was exactly correlated with that having a lower *b*-value ($r = 0.92$, $P < 0.0001$). These results confirm that the fiber tracts with a *b*-value of 4000 s/mm^2 faithfully replicated the properties of those with a lower *b*-value.

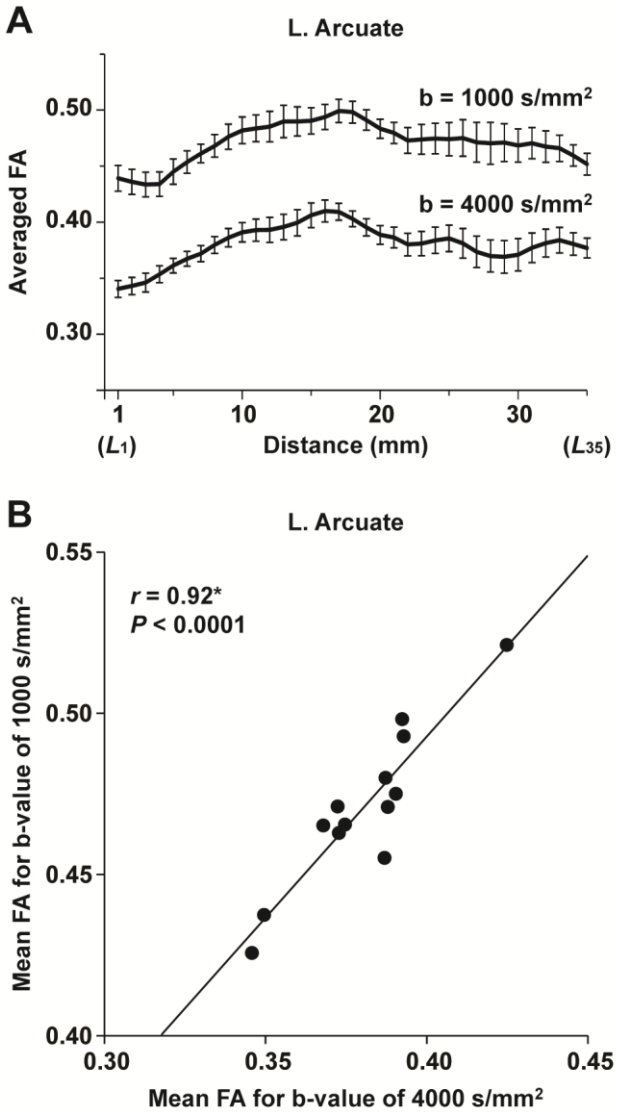

Figure 2.7. The consistency of FA for different imaging parameters. (A) At each of the L_1~L_{35} cross-sections along the left Arcuate (see Figure 2.2A), FA with a *b*-value of 4000 s/mm^2 (voxel size: 3 mm^3) or 1000 s/mm^2 (voxel size: 2 mm^3) was further averaged among 13 participants (error bars, standard errors of the mean). (B) A correlation between FA for different *b*-values. The mean FA among the voxels within L_1~L_{35} is plotted for each participant.

2.4. Discussion

The main results are summarized as follows. First, we improved the seed definition in a semi-automatic manner to reliably track the Arcuate and IFOF (Figure 2.1), and successfully identified those fiber tracts in both hemispheres of all participants. We then objectively selected their ROIs, thereby minimizing the variances in thickness among participants (Figure 2.2), and further clarified that the Arcuate was significantly thicker in the left hemisphere than in the right, while the IFOF was marginally thicker in the right hemisphere (Table 2.3). Secondly, we revealed that the mean FA in the ROI of the left Arcuate was significantly correlated with the accuracy of the Syn task (Figure 2.4A). This positive correlation was independent of gender, handedness (LQ), and the orthographic knowledge while reading sentences (the accuracy of the Spe task), as well as of age, AOE, and DOE, which were already controlled among the participants (Table 2.1). Moreover, FA in the left Arcuate was not significantly correlated with the performance of the verbal fluency task. Because the accuracy of Syn was independent of that of Spe (Figure 2.3A), the accuracy of Syn mostly represented individual syntactic abilities in L2. Thirdly, within monozygotic twin pairs, significant correlations were observed in the RTs of Syn/Spe, the accuracy of Spe, and the verbal fluency in L1, while the accuracy of Syn was marginally correlated (Figure 2.5). Indeed, the correlation coefficient for the RTs of Spe was significantly larger than that for the accuracy of Syn. Finally, the mean thickness in the ROI of the left Arcuate was highly correlated within twin pairs (Figure 2.6A). The correlation coefficient for the thickness of the left Arcuate was significantly larger than that of the left IFOF. These results suggest that general task performances, as well as the thickness of the left Arcuate, are more associated with the shared genetic/environmental factors, whereas both of mutually correlated FA in the left Arcuate and individual syntactic abilities in L2 may be less prone to these shared factors.

There have been a few previous studies in which a ROI was selected to exclude branching or curved portions of the Arcuate. In a tractography study, a ROI of a 7-mm-long

segment was chosen at a tightly bundled portion in each of the Arcuate and SLF, with some overlaps between the resultant ROIs for the two bundles (Tsang et al., 2009); those ROIs were defined in reference to the individual central sulcus, and the mean FA was obtained within each ROI. In another study, a ROI of a 10-mm-long segment was chosen in the Arcuate, where the averaged FA among participants was highest in the entire tract, instead of a curved region with lower FA (Yeatman et al., 2011). As a result of these selection procedures, the selected ROIs had similar positions in those studies, but some participants might have had lower FA with the ROIs. Moreover, it was not clear whether the ROIs were sufficiently large to represent the tract itself. Because our ROIs with uniform thickness were independent of FA, the present results would be free from any sampling bias of selecting particular regions with higher FA. In the present study, the mean FA in the ROI of the IFOF was significantly higher than that of the Arcuate in both hemispheres. The ROIs of the IFOF did not include the thinnest portion that was previously found to have a lower and more unstable FA (Yeatman et al., 2012), probably due to other intermixed tissue; such portions should be avoided in examining the relationship between FA and behavioral measurements. While FA did not show a hemispheric difference for the Arcuate or IFOF in the present study, we found that the left Arcuate was clearly thicker than the right, a finding which may be related to the left-lateralized language functions. The IFOF, on the other hand, was marginally thicker in the right hemisphere. This opposing lateralization of the Arcuate and IFOF is consistent with a previous study, where the laterality was assessed by the number of streamlines in tractography (Thiebaut de Schotten et al., 2011). The thickness in our ROIs was similar to the relative number of streamlines of each pathway, while successfully excluding the portions of branching or contamination with other tissue. Our semi-automatic methods of seed definition and ROI selection would be applicable to other pathways as well, and would thus be of use for the investigation of hemispheric or individual differences.

There were some limitations in the present methods. We minimized subjective procedures in seed and waypoint selection, but large individual variability hampered the totally automatic selection, and therefore slight manual adjustments were still necessary. This might have affected the results especially of the right Arcuate, which is known to be highly variable among individuals. However, the degree of adjustment was small in both hemispheres (mean, ≤ 5 mm) compared to the mask sizes (15–30 mm), and all tracts were inspected to confirm that they were not touching the peripheral edges of the masks (see "Materials and methods" section). Another potential concern was the thresholding used in the probabilistic tracking, in that there is no gold standard for the thresholding value. For connectivity probability values in each voxel, we used a threshold of 0.2% against individual waytotal values (i.e., values of the number of total *successful* streamlines) (Galantucci et al., 2011; Ohta et al., 2013). Some studies used a fixed percentage [e.g., 0.001% in (Griffiths et al., 2013)] against the number of all *generated* streamlines, which differed individually due to the different seed size for each participant. Given that most of the generated streamlines are rejected when reaching exclusion masks, or when going away from waypoint masks, the thresholding with successful streamlines would be more reliable. For determining an appropriate thresholding value, it is also important to consider the decrease in connectivity probability with distance from the seed mask (Morris et al., 2008). Such distance effects, as well as the concern that participants with lower waytotal values might result in a very low threshold, should be properly handled in future studies. Advances in thresholding would also contribute to examine the individual branching patterns, and to identify the anterior target regions, which might explain individual language abilities. The ROIs in the present study may depend on the gross anatomy of each tract, and further research using appropriate methods for quantifying individual variabilities in such features as branching patterns and curvature will be needed to understand the functions of language-related pathways.

The correlation coefficient between the RTs of Syn and Spe was significantly larger than that between the accuracy of the two tasks, indicating that the accuracy reflected distinct abilities required in each task, while the RTs reflected general cognitive factors, such as task difficulty and reading proficiency, common to both tasks. Moreover, short-term memory was not required in the present tasks. Although it is generally difficult to exclude the effects of semantic knowledge from any tasks using natural sentences, the ungrammatical sentences cannot be judged as incorrect by semantic information alone (see "Stimuli and tasks" section). Other remaining factors were further excluded from our partial correlation analyses between the accuracy of Syn and FA, by controlling the accuracy of Spe, because sentence reading was common to both tasks. Regarding the relationships between general cognitive abilities and L2 abilities, a previous study has reported that L2 reading proficiency was related to performances in a paired-associate test for unrelated word pairs and a phonological short-term memory task; this study used a general proficiency scale for L2 (The Interagency Language Roundtable scale) (Linck et al., 2013). Future large-scale studies using a variety of behavioral measures are necessary to reveal which network is specifically involved in syntactic, semantic, and other processes supporting language functions.

With respect to the connectivity of the dorsal pathway, some studies have indicated that the fronto-temporal connection of the left Arcuate is further separable into two parallel pathways—one terminating in the posterior superior temporal gyrus, and the other terminating in the middle temporal gyrus (Glasser and Rilling, 2008)—while other studies have suggested that there are three segments in the Arcuate/SLF: a *long* segment between the frontal and the superior/middle temporal regions, a *posterior* segment between the inferior parietal and temporal regions, and an *anterior* segment/SLF between the frontal and inferior parietal regions (Catani et al., 2012). Moreover, the SLF may be further divided into three branches (Thiebaut de Schotten et al., 2012), whose correspondence to the anterior segment/SLF remains unclear. In regard to the ventral pathway, it includes the IFOF and

uncinate fasciculus (UF) connecting the ventral frontal regions. The IFOF connects the ventral IFG (BA 45/47) and some posterior regions in the parietal, posterior temporal, and occipital cortex (Martino et al., 2010), whereas the UF connects the orbitofrontal cortex (BA 10/11/47) and anterior temporal lobe, including the temporal pole (Catani et al., 2012). The ventral pathway may also include partial connections, such as the extreme capsule (EmC), external capsule (EC), middle longitudinal fasciculus (MdLF), and inferior longitudinal fasciculus (ILF) (Makris and Pandya, 2009; de Champfleur et al., 2013), although their structural complexity, as well as the available scanning resolution, hampers the dissociation of these connections. There is thus need of a method to clarify the detailed connectivity of these pathways. One critical method would be to define seeds reliably and objectively, thereby combining the results from various imaging techniques, including diffusion MRI, functional MRI, and VBM.

The functional roles of the dorsal and ventral pathways have also been a matter of debate. A previous patient study reported that reduced FA in the left Arcuate was related to the degree of syntactic deficits in patients who suffered from primary progressive aphasia with some variants (Wilson et al., 2011), while other studies have claimed other functional roles, such as phonemic processing, lexico-semantic processing, or word learning (Glasser and Rilling, 2008; López-Barroso et al., 2013). As regards the ventral pathway, a previous functional/diffusion MRI study has suggested that the left ventral pathway, including the EmC, MdLF and ILF, is the sound-to-meaning mapping stream, focusing on regions activated for the contrast of meaningful speech vs. pseudospeech (Saur et al., 2008), and we have previously suggested that the left ventral IFG is involved in the selection and integration of semantic information (Homae et al., 2002). In our recent study of both normal and agrammatic patients, we demonstrated the existence of three syntax-related networks (Kinno et al., 2014): Network I includes the left dorsal IFG and a portion of the right dorsal pathway, and subserves syntax and its supportive system, and Network II includes a portion of the left

dorsal pathway, working as a syntax and input/output interface, while Network III includes a portion of the left ventral pathway, and is related to the syntax-semantic interaction. As demonstrated by the present study, the left Arcuate, i.e., the left dorsal pathway of Network II, has more functional significance among these syntax-related pathways, in that this pathway's FA reflected the individual syntactic abilities in L2. The left dorsal pathway would probably be involved in both bottom-up and top-down processing between lexico-semantic information and syntactic computation, consistent with previously proposed models (Hagoort, 2005; Friederici, 2012).

Several models for the functions of dorsal and ventral language pathways have been proposed. For instance, Friederici and others have proposed that the ventral pathways support syntactic phrase structure building, while the dorsal pathways, especially the one connecting the BA 44 and temporal regions, support the processing of syntactically complex sentences (Friederici, 2012). Hagoort and others have proposed another model, in which two crucial components of language systems, i.e., Memory (the mental lexicon) and Unification (the on-line assembly of lexical elements into larger structures at syntactic, semantic, and phonological levels), are represented in the temporal/inferior parietal regions and the left IFG, respectively, connected by the dorsal and ventral pathways (Hagoort, 2014). Bornkessel-Schlesewsky and others have proposed that the ventral pathways are engaged in time-independent auditory object recognition and combination, whereas the dorsal pathways are engaged in time-dependent sequence processing (Bornkessel-Schlesewsky et al., 2015). Our results support the critical role of the dorsal pathway in the light of processing argument structures of verbs in L2 acquisition. Further research with standardized tests for performance in English as a foreign language that additionally cover other domains of syntactic ability is needed to reveal how various subcomponents of syntactic processing are subserved by the dorsal and ventral pathways. Any approaches to combine the anatomical significance of these

dorsal/ventral pathways with findings from developmental and lesion studies would be useful in elucidating how particular pathways contribute to subprocesses in language functions.

According to a previous study on 9-year-old monozygotic and dizygotic twins, the contribution of non-shared environmental factors to individual variations of FA, which was averaged in the entire left SLF (probably extended to the Arcuate), was approximately 65%, while the contributions of genetic and shared environmental factors were 30% and 5%, respectively (Brouwer et al., 2010). Consistent with such contribution of non-shared environmental factors, the local FA can increase after several months of skill or vocabulary training even in adults (Scholz et al., 2009; Hosoda et al., 2013). On the other hand, it has been reported that the white matter volume was more dependent on genetic factors; the genetic contribution to the whole white matter volume was 85%, while the contribution of genetic factors to its FA was as low as 25% (Brouwer et al., 2012). There are thus at least three independent issues that should be clarified for twin studies in neuroscience: (1) shared vs. non-shared genetic/environmental factors; (2) local vs. whole brain; and (3) gray vs. white matter. In the present study, we showed that the mean thickness in the ROI of the left Arcuate was highly correlated within monozygotic twin pairs, while its FA was not. A previous MRI study on the gray matter of twins showed stronger shared genetic influences on the density (the proportion of gray matter) of the left IFG, which were more prominent than the genetic influences on the right IFG density (Thompson et al., 2001). Although twin studies on L2 acquisition are limited, a recent study with twins aged 14 has reported that 25% of the variance of L2 was due to non-shared environmental factors, where their L2 performances were evaluated by the UK National Curriculum rating system (Dale et al., 2012). Here we report that the effects of shared factors depended differentially on the L2 tasks, the accuracy, and RTs. More specifically, the correlations within twin pairs were strikingly high for RTs of both Syn and Spe, which reflected the speed of word recognition, reading, and/or decision making. Moreover, the high correlations for the accuracy/RTs of Spe, as well as those for

verbal fluency in L1, suggest that such word-level knowledge is also strongly associated with shared genetic/environmental factors, while the marginal correlation for the accuracy of Syn suggests that syntactic abilities are less prone to such shared factors.

Chapter 3.

Differential signatures of second language syntactic performance and age on the structural properties of the left dorsal pathway

3.1. Introduction

Given the large individual differences in L2, many complicated issues should be tackled in examining the neural plasticity related to L2 acquisition. In addition to participants' current age, the age of first exposure (AOE) and the duration of exposure (DOE) represent other factors in L2 acquisition (Li et al., 2014). In the present study, we controlled AOE in order to examine any group differences related to the current ages of participants (hereafter, age-related group differences), and therefore recruited students at two ages, i.e., junior (age: 13–14) and senior (age: 16–17) high-school students. Their AOE to English was 12 or 13, and they attended the same school where English classes were based on the curriculum guidelines determined by MEXT (Ministry of Education, Culture, Sports, Science and Technology). The temporal factor of L2 experience, as represented by the DOE, was also controlled and eliminated from the performances for the students at the same age. Even among such students, large individual variations in L2 performance were observable. Given that Japanese students tend to make similar mistakes in English, such as applying the null-subject (pro-drop) allowed in Japanese, we examined participants' syntactic abilities in English by a syntactic error-detection task (Syn) that we previously developed with high school teachers (Sakai et al., 2009). Orthographic knowledge in English was also examined by using a spelling error-detection task (Spe) with basically the same sets of sentences. Based on the task performances, the junior and senior students were separately divided into subgroups. As a result, the task performances of the Junior (High) and the Senior (Low) groups matched, while those of the Senior (High) group were significantly better. We then compared these three groups, with either age or L2 performance being fixed. After the group division, we identified the dorsal and ventral language-related pathways, two major routes that combine multiple language areas (Hickok and Poeppel, 2007; Friederici, 2011). The dorsal and ventral pathways correspond to the fronto-temporal segment of the arcuate fasciculus (Arcuate) and the inferior fronto-occipital fasciculus (IFOF), respectively. We identified these pathways

using the methods with diffusion MRI, which we established in Chaper 2 (Yamamoto and Sakai, 2016). We focused on two distinct structural properties, i.e., the thickness (or the volume of a tract) and FA, to examine how age and L2 performance are reflected in these properties of the dorsal and ventral pathways in each hemisphere.

Several diffusion MRI studies have suggested that the dorsal pathways mature later than the ventral pathways in both infants and children. A diffusion MRI study with tractography has reported that one of the two dorsal pathways, the one connecting the IFG and the temporal cortex, was not trackable in two-day-old newborns; in contrast, the IFOF of the ventral pathways, as well as the other dorsal pathway connecting the premotor cortex and the temporal cortex, was already present (Perani et al., 2011). Another recent study has reported that the maturation of the ventral pathways in infants aged 6–22 weeks was more advanced than that of the dorsal pathways in terms of microstructural properties including FA, although the difference between the dorsal and ventral pathways decreased with increasing weeks of age (Dubois et al., 2016). In a study with seven-year-old children, both the dorsal and ventral pathways were trackable, but FA of these pathways was significantly lower than that of adults (Brauer et al., 2013), suggesting that the language-related pathways were not fully mature at this age. Regarding the adolescent ages, it has been suggested that the dorsal pathway was still under development, and that the ventral pathway showed less prominent development (Lebel and Beaulieu, 2011), but it remains unclear how these structural developments can be correlated with behavioral measurements. The relatively slow development of the dorsal pathway suggests the interesting possibility that this pathway may be more strongly influenced by language experiences.

We have performed the following functional and structural studies on adolescent students. In our previous functional MRI study, we identified cortical regions involved in syntactic processing using the same English tasks as in the present study, and showed a positive correlation between the individual activations of the left IFG and the performance

accuracy of the Syn task (Sakai et al., 2009). Our VBM study further clarified that individual leftward lateralization of a single region in the IFG was also correlated with the accuracy of Syn (Nauchi and Sakai, 2009). Moreover, in Chapter 2, we reported that individual FA in the left Arcuate was correlated with the accuracy of Syn (Yamamoto and Sakai, 2016), suggesting the importance of the left dorsal network including the IFG and the Arcuate in syntactic processing. More specifically, among the Arcuate and IFOF in both hemispheres, we revealed that FA in the left Arcuate alone showed a significant positive correlation with the accuracy of the Syn task, but not with the accuracy of Spe. Further, within monozygotic twin pairs, neither the accuracy of Syn nor FA in the left Arcuate were significantly correlated between the twins, in spite of the high inter-twin correlation for the thickness of the left Arcuate. Given these results, syntactic abilities in L2 and FA in the left Arcuate may thus have been sensitive to non-shared environmental factors by which the twins were individually affected, while the thickness was dependent on shared genetic/environmental factors. Based on these points, we made two further hypotheses. First, FA in the left Arcuate should show performance-related group differences that would be observable even among the students with the same DOE, because larger variances within the monozygotic twin pairs were observed for FA than for the thickness. Second, the thickness of the left Arcuate should show age-related group differences, as a similar thickness was observed for monozygotic twin pairs. We tested these hypotheses, using semi-automatic methods of defining seeds for tractography and selecting ROIs, which we developed in Chapter 2 (Yamamoto and Sakai, 2016). In this previous study, we showed that the thickness in a one-dimensional ROI of the left Arcuate was clearly larger than that of the right Arcuate; such laterality was evident neither for the thickness of the IFOF, nor for FA in the Arcuate or IFOF. In the present study, we confirmed that the leftward laterality of the Arcuate was consistent among the three groups. We further examined the correlation with individual accuracy of Syn/Spe to clarify which particular aspects of L2 abilities were related to the structural properties that showed

any performance-related group differences. Our results should thus help to elucidate the neuroanatomical basis of language acquisition after the sensitive period.

3.2. Materials and methods

3.2.1. Participants

We recruited junior high-school students in their second academic year, as well as senior high-school students in their fifth academic year, at the Secondary Education School attached to the Faculty of Education of the University of Tokyo, to which twins are preferentially admitted for educational research. The participants performed the Syn, Spe, and verbal fluency tasks as in Chapter 2. We set the following accumulative inclusion criteria: (i) right-handedness, i.e., a positive LQ, as assessed by the Edinburgh inventory (Oldfield, 1971), (ii) no history of neurological or psychiatric diseases, (iii) native Japanese speakers whose AOE to English in formal education was 12 or 13 years old (a condition met by the majority of students in this school), and (iv) RTs for Spe within the presentation time (6400 ms) for more than 90% of the trials. For each twin pair who met these criteria, the one exhibiting the higher score for the Spe task was entered into the analysis to avoid a double count of twins with potentially similar characteristics. As regards the first criterion, eight junior and 11 senior students with a negative LQ or those with a potential history of change in handedness were dropped, resulting in a relatively strong right-handed population (LQ > 35). One participant each for the second and fourth criterion, and two participants for the third criterion, were dropped.

For the senior high-school students, we employed three additional criteria: (v) an accuracy of Spe higher than 65% (i.e., mean – 1.5 SD), (vi) no worse outliers in each task (i.e., higher than the mean – 2 SD for the accuracy, and shorter than the mean + 2 SD for the RTs), and (vii) shorter RTs for easier Spe than for Syn. Given that most Japanese students learn an alphabetic writing system at the age of 12−13, and that the senior high-school

students had been studying English for about four years, it was necessary to set the fifth criterion in order to exclude the potential effects of poor reading abilities and precisely assess individual syntactic abilities; five senior high-school students were dropped for this reason. Two students were dropped because they did not meet the sixth criterion, and two more students because they did not meet the seventh criterion. As a result, we enrolled a total of 39 junior and 38 senior high-school students.

We divided the junior high-school students into two groups: a group of 14 students (the Junior (High) group) who scored higher than 65% in Spe (this was identical to the fifth criterion employed for senior high-school students), and a group of 25 students (the Junior (Low) group) with scores lower than 65% in Spe, which were too low to assess their L2 abilities and related structures further. In regard to the senior high-school students, we first divided the 38 students into two groups using K-means cluster analysis (R software, https://www.r-project.org/) on the accuracy of Syn and Spe tasks. This analysis yielded a group of 15 participants with higher L2 abilities (the Senior (High) group), as well as a group of 23 participants with lower L2 abilities. Because the performances of the group with lower L2 abilities did not match those of the Junior (High) group for the accuracy and RTs of Spe (one-sided *t*-tests, $P < 0.05$), we further performed a hierarchical cluster analysis using Ward's method (R software) based on the accuracy and RTs of Spe for the 23 students with lower L2 abilities. As a result, we divided these students into two groups: a group of 5 students with the higher accuracy and shorter RTs of Spe (the Senior (Middle) group), and a group of 18 other students (the Senior (Low) group). The accuracy/RTs of both tasks for the Senior (Low) group matched those for the Junior (High) group (one-sided *t*-tests, $P > 0.05$). Because our purpose was to examine the differences between the performance-matched groups with different ages, as well as between age-matched groups with different L2 abilities, we focused on the Junior (High), Senior (Low), and Senior (High) groups, dropping the other groups. All participants in the Senior (High) group, as well as seven participants in the Senior

(Low) group, were included in Chapter 2, in which we analyzed 26 participants with an accuracy of Spe ≥ 80% (Experiment I) (Yamamoto and Sakai, 2016); moreover, these 26 former participants were all included in the group of 38 senior high-school students mentioned above. Participants in the Junior (High) group were newly recruited for the present study. All participants in the three groups were scanned with the same diffusion MRI protocol. Demographic details of participants in these three groups are shown in Table 3.1: age, AOE to English, DOE to English, and LQ. Written informed consent was obtained from all participants as well as from their parents/guardians. The study was approved by the Secondary Education School and by the Institutional review board of the University of Tokyo, Komaba.

Table 3.1. Demographic details of the participants in each group.

	N	Age	AOE	DOE	LQ	vf
Junior (High)	14	14 ± 0.4	12.6 ± 0.3	1.5 ± 0.3	84 ± 16	22 ± 3.6
	(7 females)	(13–14)	(12–13)	(1.3–1.9)	(54–100)	(16–31)
Senior (Low)	18	17 ± 0.3	12.7 ± 0.3	4.4 ± 0.2	86 ± 19	23 ± 5.4
	(10 females)	(16–17)	(12–13)	(4.3–4.8)	(38–100)	(15–33)
Senior (High)	15	17 ± 0.4	12.6 ± 0.3	4.5 ± 0.2	83 ± 15	27 ± 6.0
	(9 females)	(16–17)	(12–13)	(4.3–4.8)	(56–100)	(17–36)

The number of participants (N), age, age of first exposure (AOE) to English, duration of exposure (DOE) to English, laterality quotient (LQ), and the number of words produced in the verbal fluency task (vf), are shown as the means ± SD. Ranges are shown in brackets.

3.2.2. Identification of the Arcuate and IFOF

MR imaging acquisition was conducted with the same parameters as in Experiment I in Chapter 2. The Arcuate and IFOF were identified in each individual hemisphere (see "Materials and methods" section in Chapter 2). In the present study with group comparisons, all the tracts were normalized and analyzed after the tracking in individual brains; in Chapter 2, tracts were analyzed in the native space to focus on individual variabilities of structural indices (Yamamoto and Sakai, 2016). Using the affine and nonlinear transformation with

FLIRT and FNIRT, all the tracts were spatially normalized to the MNI space, and were binarized with the fslmaths function of FSL. The binarized tracts were overlaid across participants to produce population probability maps for each pathway, in which voxel values represent the number of participants. The population probability maps with thresholding (at least half of the participants) were smoothed and shown using MRIcroN software (http://people.cas.sc.edu/rorden/mricron/index.html).

3.2.3. ROI selection

We basically followed the ROI selection procedures in Chapter 2 (Yamamoto and Sakai, 2016); we determined one-dimensional ROIs (in an antero-posterior direction) at the portion with the most uniform *thickness*, and showed that the appropriate length, i.e., *ROI size*, minimizing the individual variances of thickness was 20 and 15 mm for the Arcuate and IFOF, respectively. The thickness of the fiber tract was defined as the number of voxels (voxel size, 1 mm^2) at a certain section of the tracked fibers. Here we chose coronal sections, because the portions without branching in the Arcuate or IFOF were nearly horizontal. We measured the thickness along the Arcuate in each hemisphere for the length of 35 mm ($y = -40 \sim -6$; the candidate region for ROI), excluding the branching or curved portions. Across a tract segment of 20 mm, we calculated the SD of the tract's thickness, and slid the segment within these candidate regions. At each position, the averaged SD was obtained among all the participants. The segment with a minimal averaged SD, i.e., with the most uniform thickness across participants, was selected as an ROI for the Arcuate in each hemisphere. Using these procedures, we objectively selected ROIs at the same position for all the participants, thereby minimizing individual variabilities in the ROI selection.

In regard to the IFOF, we set a candidate region with a length of 70 mm and selected 15-mm-long ROIs, in accordance with our previous study. We measured the thickness along the IFOF in each hemisphere for $y = -75 \sim -6$, excluding the branching or curved portions.

Across a tract segment of 15 mm, we calculated the SD of the tract's thickness, and slid the segment within these candidate regions. At each position, the averaged SD was obtained among all the participants. The segment with a minimal averaged SD, i.e., with the most uniform thickness across participants, was selected as an ROI for the IFOF in each hemisphere.

We next examined the correlation with individual accuracy of Syn/Spe to clarify which aspects of L2 abilities were related to the structural property that showed any performance-related group differences. For the analyses within a single group, we aimed to examine individual variances in a pathway, thereby employing individually selected ROIs on the normalized tracts. We selected ROIs where the thickness was most uniform for each tract in the MNI space. As described above, we measured the thickness along the Arcuate in each hemisphere for the candidate region with a length of 35 mm ($y = -40 \sim -6$). Across a tract segment of 20 mm, we calculated the SD of the tract's thickness, and slid the segment within these candidate regions. Among these segments, the segment with a minimal SD, i.e., with the most uniform thickness, was selected as an ROI for each tract of the Arcuate.

3.2.4. FA analyses

As described in the above sections, ROIs were determined on the basis of the connectivity probability and resultant thickness. After selecting ROIs, voxels with FA values of 0.2 or higher were thresholded only in FA analyses, and the resultant FA maps were normalized using the affine and nonlinear transformation with FLIRT and FNIRT. Thresholding has been widely used in previous studies to avoid the uncertainty of anisotropy measurements in the peripheral regions of fibers. We calculated the mean FA within each ROI, and examined the group differences. For the partial correlation analyses, we used the ppcor (partial and semi-partial correlation) package (https://cran.r-project.org/web/packages/ppcor/) in R software (http://www.r-project.org/).

3.3. Results

3.3.1. Behavioral data

Behavioral data of the Syn and Spe tasks in the Junior (High), Senior (Low), and Senior (High) groups are shown in Figure 3.1. Regarding the accuracy (Figure 3.1A), a two-way rANOVA [group × task (Syn, Spe)] showed significant main effects of group [$F(2, 44) = 47$, $P < 0.0001$] and task [$F(1, 44) = 169$, $P < 0.0001$] without interaction [$F(2, 44) = 0.4$, $P = 0.6$]. Regarding RTs (Figure 3.1B), an rANOVA [group × task (Syn, Spe)] showed significant main effects of group [$F(2, 44) = 6.3$, $P < 0.005$] and task [$F(1, 44) = 116$, $P < 0.0001$] with a significant interaction [$F(2, 44) = 4.2$, $P < 0.05$]. According to the *t*-tests (Bonferroni-corrected for three comparison pairs, significance level at $\alpha = 0.017$), we confirmed that the Senior (High) group had significantly higher accuracy and shorter RTs than the other two groups in both tasks, indicating that the Senior (High) group had better syntactic abilities and word knowledge than the other two groups.

If a behavioral parameter reflected general cognitive factors, such as task difficulty and reading proficiency, which were commonly involved in the Syn and Spe tasks with some differences in degree, then that parameter would be expected to show a correlation between Syn and Spe among the participants, when considering their individual differences in these factors (Yamamoto and Sakai, 2016). For example, a slower reader would show longer RTs for both Syn and Spe irrespective of the tasks. Therefore, as a contraposition, if another behavioral parameter did not show such a correlation between Syn and Spe among the participants, that parameter would reflect, at very least, the distinct factors required in each task. We thus examined the relationships between the performances of the two tasks. The accuracy of Syn and Spe did not show a significant correlation ($r = 0.27$, $P = 0.07$; Figure 3.1C), suggesting that the accuracy mainly reflected distinct abilities for each task. On the other hand, the RTs of Syn and Spe showed a highly positive correlation ($r = 0.82$, $P < 0.0001$; Figure 3.1D). Indeed, the RTs of Syn and Spe showed a significant positive

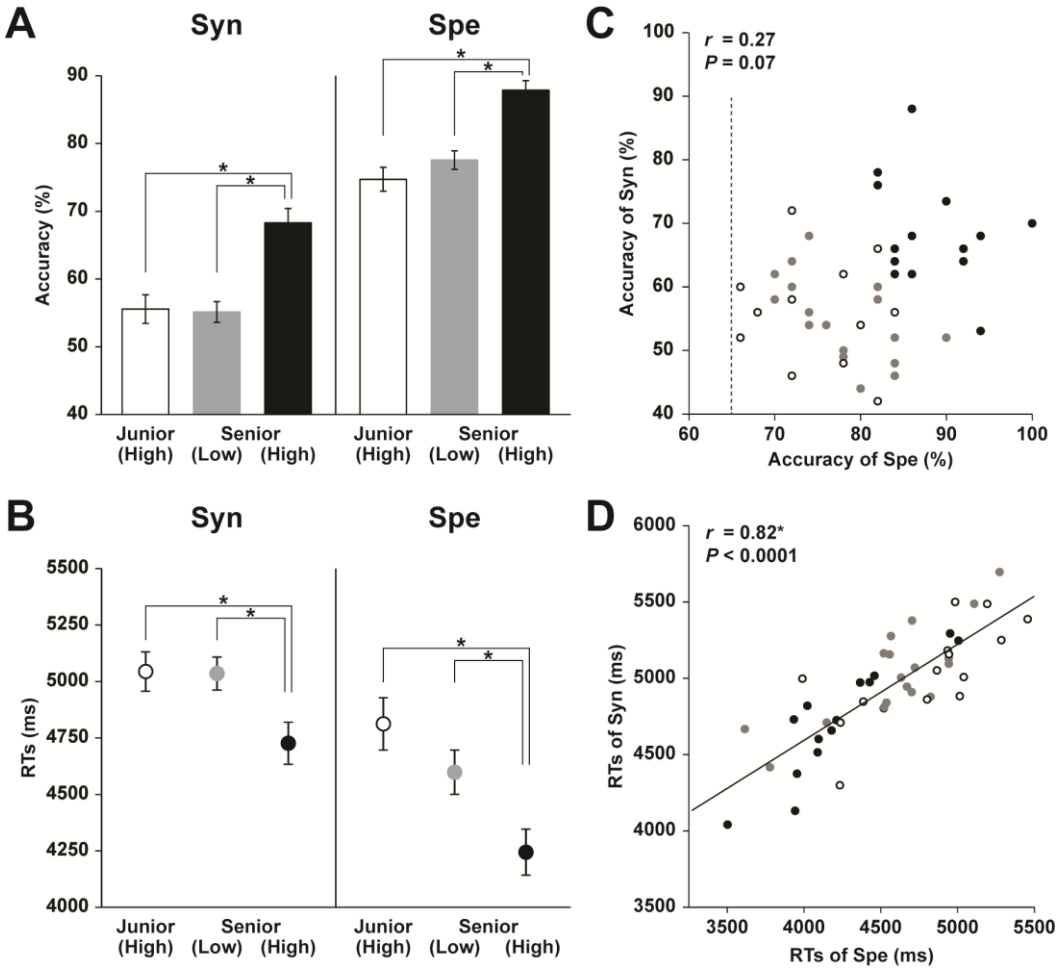

Figure 3.1. Behavioral data of Syn and Spe tasks for the three groups. (A) The accuracy of Syn and Spe tasks for the three groups of high school students, Junior (High), Senior (Low), and Senior (High), are indicated with white, gray, and black bars, respectively. Error bars indicate the standard error of the mean (SEM) for the participants, and asterisks denote the significant differences (Bonferroni-corrected). (B) The reaction times (RTs) of the Syn and Spe tasks for the three groups are indicated by dots corresponding to the white, gray, and black shades in (A). The Senior (High) group showed significantly higher L2 abilities, i.e., higher accuracy and shorter RTs, while the Junior (High) and Senior (Low) groups showed no significant difference in the accuracy or RTs in either task. (C, D) Independence of the accuracy of the Syn and Spe tasks. Individual behavioral parameters were plotted to compare the two tasks, as shown by the plotted dots, which correspond to the shades in (A). The dotted line indicates an inclusion criterion of Spe. The accuracy of Syn and Spe showed no significant correlation (C), while the RTs of Syn and Spe were highly correlated (D).

correlation in all three of the groups ($r > 0.7$, $P < 0.005$). These results indicate that the RTs reflected general cognitive processes common to both tasks. We further observed that the correlation coefficient between RTs was significantly larger than that between the accuracy ($Z = 4.4$, $P < 0.005$) by using the cocor (comparing correlations) package in R software,

confirming the different natures of the accuracy and RTs. These results indicate that the *accuracy* of each task for the participants reflected individual abilities required in each task. Once a structural property which showed any group difference related to L2 acquisition was found, we further examined its correlation with the accuracy of Syn or Spe to reveal which ability was dominantly reflected.

We also examined the verbal fluency data in L1, and found a consistent trend among groups with the Syn and Spe performances. Behavioral data of the verbal fluency task are shown in Table 3.1. A one-way ANOVA showed a significant effect of group [$F(2, 44) = 4.0$, $P < 0.05$]. The Senior (High) group produced significantly larger numbers of words than the Junior (High) group ($t(23) = 2.8$, $P < 0.005$), surviving Bonferroni correction for the three comparison pairs (significance level at $\alpha = 0.017$). The Senior (High) group produced larger numbers of words than the Senior (Low) group ($t(31) = 2.0$, $P = 0.03$). For the performance-matched groups, i.e., the Junior (High) group and the Senior (Low) group, there was no significant difference ($t(30) = 0.8$, $P = 0.2$). These results suggest that similar group differences were found for the L1 and L2 tasks we tested.

3.3.2. Group differences along the tract for the Arcuate

The Arcuate and IFOF were successfully tracked in both hemispheres for every participant, and the tracking was basically similar among groups. In all groups, the Arcuate connected the frontal and temporal regions with a similar curvature, and the IFOF connected the frontal and occipital/temporal regions with a narrower portion near the external capsule (Figure 3.2). While these overall characteristics were the same for the left and right hemispheres, the Arcuate was thicker in the left than in the right hemisphere. From both lateral and top views, the Arcuate was consistently thicker and extended more anteriorly in the left hemisphere. The Arcuate of the Senior (High) and Senior (Low) groups was thicker than that of the Junior

(High) group, which was evident from both views. Such group differences were not observed for the IFOF in either hemisphere.

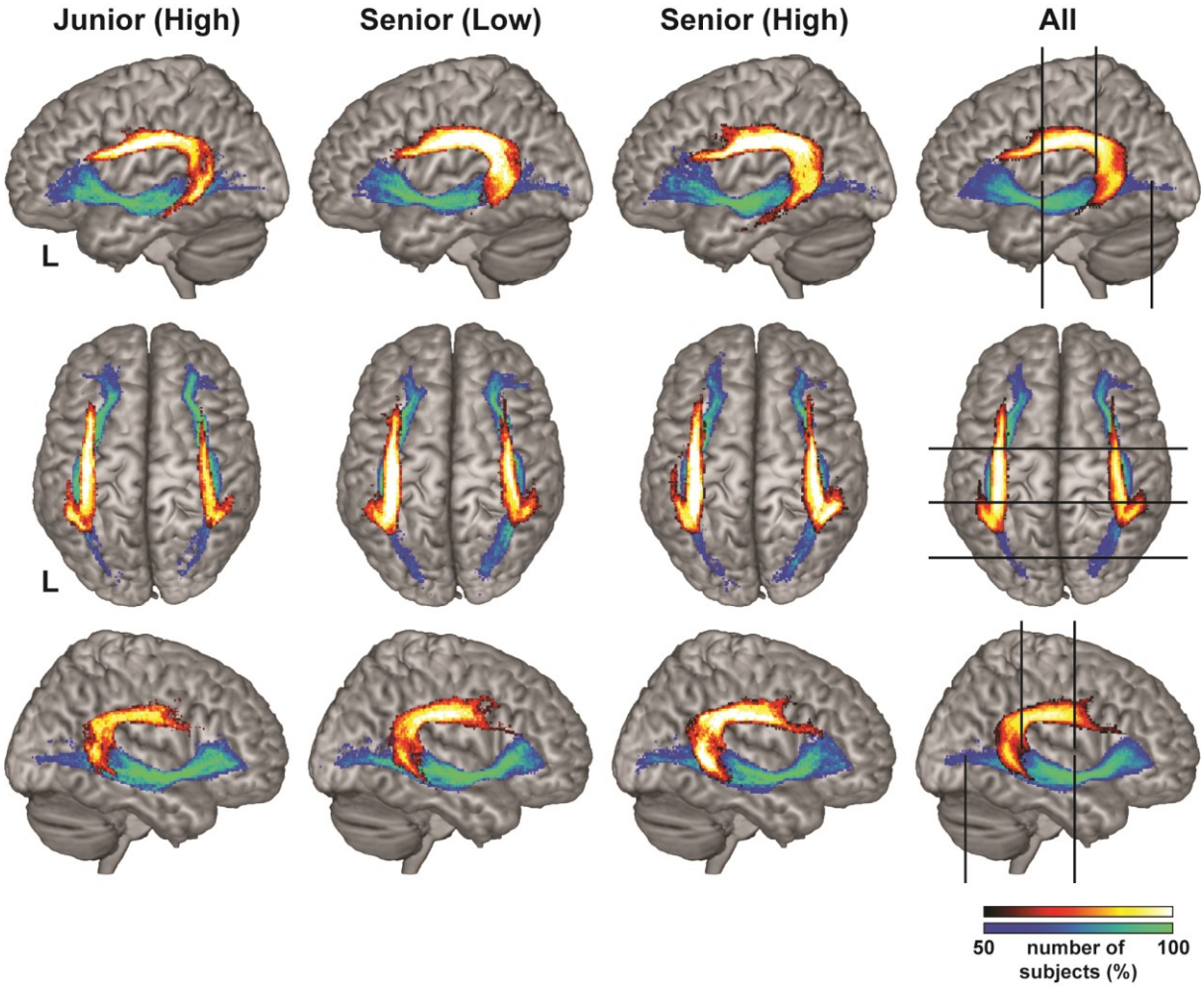

Figure 3.2. The Arcuate and IFOF reconstructed with diffusion MRI. The probability maps of population for the Arcuate and IFOF in the MNI space. The pathways were thresholded to show the results of the tracking present in more than half of the participants in each group. The color scales denote the number of participants (%): black-red to yellow-white for the Arcuate, and blue to green for the IFOF. Note that the Arcuate in the Senior (High) and Senior (Low) groups was clearly thicker than that in the Junior (High) group, which was evident from both left lateral (top row) and top (middle row) views. Similar tendency was observed in the right hemisphere from the right lateral (bottom row) view. To examine the overall profile while excluding highly variable regions among individuals, we first set candidate regions for selecting regions of interest (ROIs): a region of 35 mm for the Arcuate ($y = -40 \sim -6$), and a region of 70 mm for the IFOF ($y = -75 \sim -6$). The right-most panel shows the candidate regions, which are bounded by the pairs of black lines. We further selected ROIs at the portion with the most uniform thickness in each pathway to quantify the Arcuate and IFOF (see "Materials and methods" section), as shown in Figure 3.3.

To examine the overall profile while excluding highly variable regions among individuals, we first set the candidate regions for ROI selection: the region of 35 mm for the Arcuate (y = –40 ~ –6), and the region of 70 mm for the IFOF (y = –75 ~ –6) (Figure 3.2). As regards the Arcuate, the candidate region was selected where the pathway was relatively straight, excluding the curved or branching portions. The candidate region for the IFOF was also selected where the pathway was straight, excluding the narrower portion near the external capsule, resulting in a longer candidate region than for the Arcuate. The thickness of the Arcuate in the Senior (High) and Senior (Low) groups was significantly larger than that of the Junior (High) group throughout the candidate region in both hemispheres, as indicated by the non-overlapping error bars (mean ± SEM) (Figure 3.3A). In contrast, for the thickness of the IFOF, no clear group difference was observed in either hemisphere. For most of the

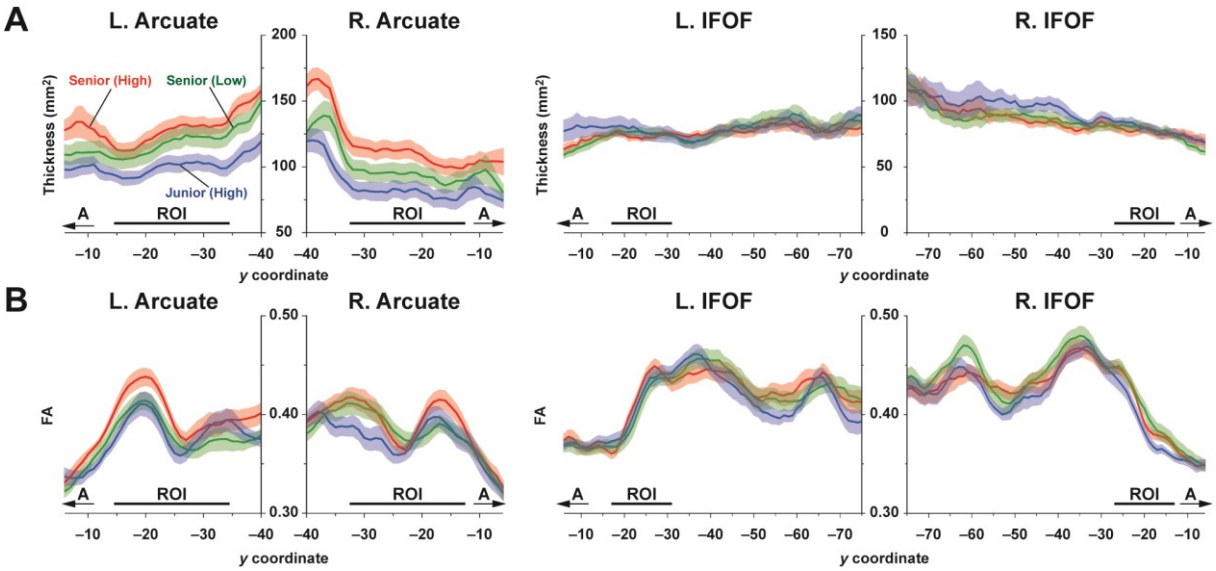

Figure 3.3. Profiles of the thickness/FA of the dorsal and ventral pathways in each group. (A) The profiles of thickness for the Arcuate and IFOF. (B) The profiles of FA for the Arcuate and IFOF. The thickness and FA were averaged among the participants in each of the Junior (High), Senior (Low), and Senior (High) groups, as shown in blue, green, and red, respectively. Data are shown for the range of y = –40 ~ –6 for the Arcuate and y = –75 ~ –6 for the IFOF (bounded by the pairs of black lines in Figure 3.2) in the MNI space. Note that the thickness of the Arcuate in the Senior (High) group is larger than that in the Junior (High) group in both hemispheres. Moreover, FA in the left Arcuate in the Senior (High) group is higher than those in the other two groups. No clear group difference was found for the IFOF in either hemisphere. The SEMs in each group are shown as shaded bands in each color. The positions of ROIs are indicated by the black lines above the axes. A, anterior.

candidate regions in the Arcuate and IFOF, the thickness was basically uniform in both hemispheres. We selected one-dimensional ROIs, where the thickness was most uniform (see "Materials and methods" section). Because the thickness is independent of FA, these ROIs were free from any sampling bias of selecting particular regions with higher FA. We also plotted FA in the same regions in each hemisphere, and found that FA of the left Arcuate in the Senior (High) group was higher than those of the other two groups, especially in the anterior regions, as indicated by the non-overlapping error bars (Figure 3.3B). In regard to the IFOF, however, no clear group difference was observed. In these ROIs, FA in the Arcuate showed some modulations, while FA in the IFOF showed relatively sharp antero-posterior changes. Nevertheless, the overall tendency of thickness or FA was similar among the three groups throughout the candidate regions in both hemispheres.

3.3.3. Distinct group differences in the structural properties of the left Arcuate

For all the participants, the ROIs were placed at the same position of each pathway in the MNI space. Using these ROIs, we first confirmed the leftward laterality of the thickness of the Arcuate among the three groups (see "Introduction" section). A two-way rANOVA [group × hemisphere (left, right)] indicated significant main effects of group [$F(2, 44) = 4.9$, $P = 0.01$] and hemisphere [$F(1, 44) = 27$, $P < 0.001$], without an interaction [$F(2, 44) = 0.3$, $P = 0.7$] (Figure 3.4A). Indeed, in all three of the groups, the thickness of the left Arcuate was significantly larger than that of the right Arcuate ($P \leq 0.02$) (one-sided *t*-tests). As regards the group differences, the thickness of the left Arcuate for the Senior (High) group was significantly larger than that of the Junior (High) group ($t(27) = 3.1$, $P = 0.002$) (one-sided *t*-tests), surviving Bonferroni correction for the three comparison pairs (significance level at $\alpha = 0.017$). Moreover, the thickness for the Senior (Low) group was larger than that of the Junior (High) group ($t(30) = 1.8$, $P = 0.04$). There was no significant difference between the Senior (High) and Senior (Low) groups ($t(31) = 0.86$, $P = 2.0$). Regarding the thickness in the

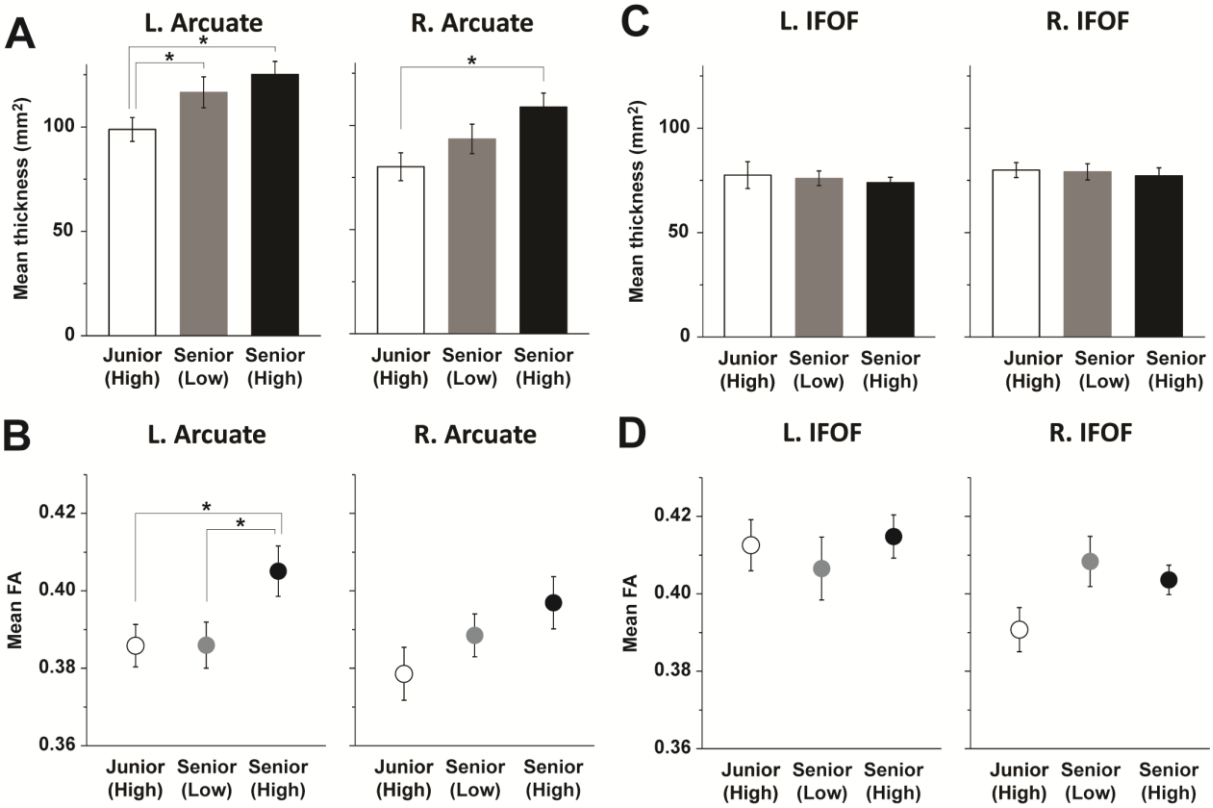

Figure 3.4. Group differences in the structural properties of the left Arcuate. (A) Age-related group differences on the thickness of the left dorsal pathway. (B) Performance-related group differences on FA in the left dorsal pathway. (C, D) The ventral pathway without group differences. The Junior (High), Senior (Low), and Senior (High) groups are shown in white, gray, and black bars/dots, respectively. Error bars indicate the SEM for the participants, and asterisks denote statistical differences at $P < 0.05$.

ROI of the right Arcuate, the thickness for the Senior (High) group was significantly larger than that for the Junior (High) group ($t(27) = 3.0$, $P = 0.003$). There was no significant difference between the Senior (Low) group and Junior (High) group ($t(30) = 1.3$, $P = 0.10$) or between the Senior (High) and Senior (Low) group ($t(31) = 1.6$, $P = 0.06$).

Another structural property of the Arcuate, i.e., mean FA in the ROI, was clearly higher in the Senior (High) group than in the other two groups in the left hemisphere (Figure 3.4B). A one-way ANOVA for FA in the left Arcuate indicated a significant main effect of group [$F(2, 44) = 3.3$, $P < 0.05$]. More specifically, FA in the left Arcuate for the Senior (High) group was significantly higher than that for the Junior (High) group ($t(27) = 2.2$, $P = 0.016$), surviving Bonferroni correction for the three comparison pairs. Moreover, FA in the

left Arcuate for the Senior (High) group was higher than that for the Senior (Low) group ($t(31) = 2.2$, $P = 0.019$). There was no significant difference between the Senior (Low) and Junior (High) groups ($t(30) = 0.02$, $P = 0.5$). Regarding FA in the ROI of the right Arcuate, a one-way ANOVA did not show a significant main effect of group [$F(2, 44) = 2.0$, $P = 0.15$].

In regard to the IFOF, neither the mean thickness nor mean FA in the ROI showed a group difference (Figure 3.4C, D). A one-way ANOVA for the thickness did not show a significant main effect of group in either hemisphere [$F(2, 44) < 1.7$, $P > 0.8$]. In addition, a one-way ANOVA for FA in the IFOF did not show a significant main effect of group in either hemisphere [$F(2, 44) < 2.6$, $P > 0.05$]. Major structural properties, i.e., the thickness and FA, in the ROI of IFOF in both hemispheres were similar among the three groups.

3.3.4. FA in the left Arcuate was selectively correlated with the accuracy of Syn

For FA in the left Arcuate, which showed group differences between the Senior (High) and the other two groups (see Figure 3.4B), we examined what aspect of L2 abilities was actually related to FA. We performed the following analyses for the Senior (High) group, whose task performances were higher and thus most reliable for dissociating the linguistic abilities required by Syn or Spe. ROIs were selected for each participant as described in the Materials and methods section. We performed partial correlation analyses between the standardized accuracy of Syn and the standardized FA in the left Arcuate, removing the effects of the accuracy of Spe, LQ, and gender. Regarding the accuracy of Syn, we found a significant correlation with FA in the left Arcuate ($r = 0.61$, $P = 0.03$) (Figure 3.5A). In contrast, the accuracy of Spe was not significantly correlated with FA in the left Arcuate ($r = 0.40$, $P = 0.2$), according to the partial correlation analysis removing the effects of the accuracy of Syn, LQ, and gender (Figure 3.5B). In addition, no significant correlation was found between FA in the left Arcuate and verbal fluency in L1 ($r = 0.24$, $P > 0.4$) in the partial correlation analysis removing the effects of LQ and gender. These results indicate that increased FA in

the left Arcuate for the Senior (High) group was related mainly to the enhanced syntactic abilities in L2.

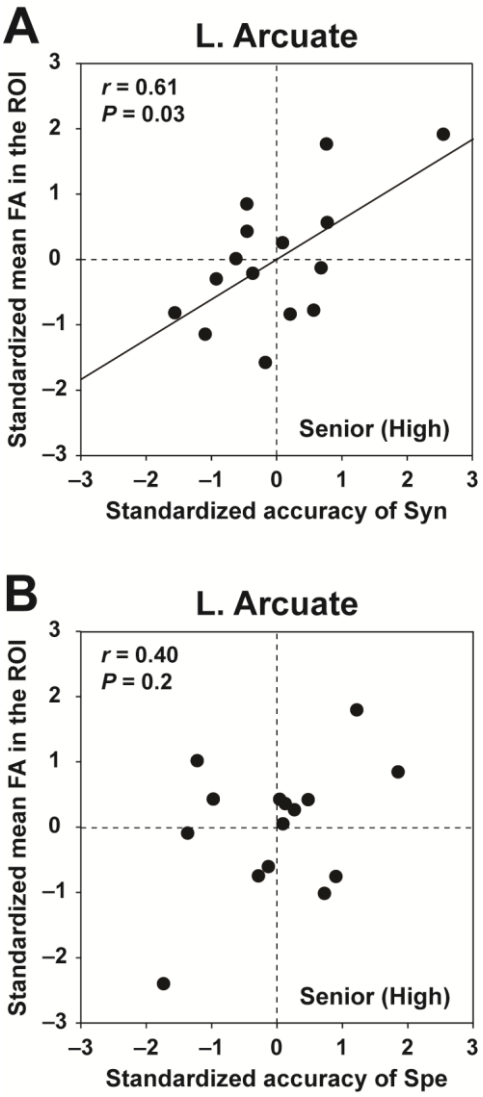

Figure 3.5. FA in the left Arcuate was selectively correlated with the accuracy of Syn. (A) The scatter plots of the standardized accuracy of Syn and the standardized mean FA in the left Arcuate. The ROIs in the left Arcuate were determined for each participant in the Senior (High) group (see "Materials and methods" section), whose FA showed performance-related group differences (Figure 3.4B). The effects of the standardized accuracy of Spe, gender, and laterality quotient (LQ) were removed. (B) The scatter plots of the standardized accuracy of Spe and the standardized mean FA in the left Arcuate. The effects of the standardized accuracy of Syn, gender, and LQ were removed. FA in the left Arcuate was correlated with the accuracy of Syn, but not with the accuracy of Spe.

3.4. Discussion

In the three groups of high school students, we obtained the following results. First, performance-related group differences were found only for FA in the left Arcuate. More specifically, the Senior (High) group, who had higher L2 abilities, showed higher FA in the left Arcuate than the Senior (Low) and Junior (High) groups (Figure 3.3B, 3.4B). Moreover, the mean FA in the ROI of the left Arcuate of the Senior (High) group was significantly correlated with the accuracy of Syn (Figure 3.5A), but not with the accuracy of Spe (Figure 3.5B) or with verbal fluency in L1, indicating that increased FA in the left Arcuate was related mainly to the enhanced syntactic abilities in L2. Secondly, age-related group differences were found for the thickness of the left Arcuate. The thickness for the Senior (High) and Senior (Low) groups was larger than that for the Junior (High) group (Figure 3.3A, 3.4A), indicating that the left Arcuate was still developing in adolescence. Thirdly, these differential performance-related and age-related signatures were evident on the left Arcuate alone, in contrast to the right Arcuate, which showed a weak indication of either signature (Figure 3.3A, 3.4A), and to the bilateral IFOF, which lacked either signature (Figure 3.3, 3.4CD). To the best of our knowledge, our study is the first to report that plasticity during the adolescent years was markedly different between the dorsal and ventral language-related pathways, the former of which was related to syntactic abilities. In summary, we showed that the left dorsal pathway, which has been reported to be more immature than the ventral pathway from early infancy, continued to grow thicker with increasing age at least until the adolescent years. Further, by showing the group difference between the Senior (High) and Senior (Low) groups, whose DOEs were the same, we revealed that performance differences were reflected in the FA in the left dorsal pathway, independent of age and the duration of experience. These results indicate that the left dorsal

pathway serves as the neuroanatomical basis for the specific and most fundamental abilities of syntax.

Within the same part of the left dorsal pathway, we observed the dissociated performance-related and age-related group differences on FA and the thickness, respectively. These results provide new insights into the plasticity of these two distinct structural properties in the left Arcuate. FA and the volume of the Arcuate have been reported to change during the adolescent years (Lebel and Beaulieu, 2011), but it has remained unclear whether or not these properties change in accordance with the development of specific abilities. The increased FA in accordance with certain learning or L2 exposure has been reported also in other pathways. For instance, an increase in FA in the genu of the corpus callosum was correlated with the overall L2 performances after a 9-month intensive course of spoken and written Chinese (Schlegel et al., 2012). Moreover, an increase in FA in the white matter underlying the right IFG was correlated with vocabulary test scores after a 16-week period of vocabulary training in English (Hosoda et al., 2013). In Chapter 2, focusing on students with better orthographic knowledge (80% or higher accuracy in the Spe task), we showed that individual differences in the accuracy of Syn were reflected in FA in the left Arcuate (Yamamoto and Sakai, 2016). In the present study, analyzing students with a wider range of L2 abilities and ages, we demonstrated that FA in the left Arcuate was higher in students with higher L2 abilities than in students with lower abilities at the same age, whose FA was not significantly different from performance-matched younger students. Further, we showed that FA in the left Arcuate of the Senior (High) group was positively correlated with the accuracy of the syntactic task, but not with the accuracy of the spelling task. These results indicate that the higher FA in the left dorsal pathway was mainly related to the higher syntactic abilities in L2. Future studies will be needed to elucidate how the development in other aspects of linguistic abilities is related to the plasticity in the left dorsal pathway.

Another important issue is whether non-cognitive abilities are also related to the development of the language-related structures. A previous study on L2 learning in infants indicated the impact of social interaction, by showing that infants who received "live-person" sessions on Mandarin performed significantly better on Mandarin phonetic perception tests, compared to infants who received the identical information via television or audiotape but showed no learning effects (Kuhl et al., 2003). These social factors, as well as other non-cognitive factors such as motivation and self-confidence, may have impacts on L2 acquisition in adolescent students and adults as well (Dörnyei, 2003; Gardner, 2010). One of the directions of future studies would be to investigate how such individual traits affect one's acquisition experience, which may further modify functional and structural networks. It would also be important to dissociate the effects of non-cognitive abilities on language-related networks from those of cognitive abilities on the left dorsal pathway, whose critical involvement in linguistic information processing has been shown by previous studies (Ohta et al., 2013; Yamamoto and Sakai, 2016), and confirmed in the present study.

Here we showed that the thickness of the left Arcuate was larger in the Senior (High) and Senior (Low) groups than in the Junior (High) group, indicating age-related group differences. This macro-structural property of the left Arcuate had plasticity associated with age, which may reflect biological maturation, as well as common and specific experiences that students underwent during the ages of 13–17. A recent longitudinal study reported that children's structural connectivity, which was obtained before they learned to read (i.e., before age 5), predicted the spatial profile of functional activations in the visual word form area (VWFA) after they learned to read (i.e., after age 8), suggesting that connectivity precedes the functional development (Saygin et al., 2016). These results raise an interesting hypothesis that experience interacts with the micro-structural development of a pathway whose structural connectivity has already developed.

In addition to the fronto-temporal segment examined here, the arcuate fasciculus is also composed of the fronto-parietal and temporo-parietal segments (Catani et al., 2005). Thiebaut de Schotten et al. (2014) have suggested that the temporo-parietal segment has plasticity associated with learning even in adults, based on their finding that its FA was higher in a group of ex-illiterates, who lacked access to schools during childhood for social reasons and learned to read during adulthood, in contrast to a group of illiterates, who never learned to read. They also showed that FA was correlated with activations in the two regions, i.e., the left VWFA and planum temporale, connected to the angular/supramarginal gyri by this segment. As the authors of this study discussed in their subsequent paper (Dehaene et al., 2015), these results might have been influenced by multiple factors, including the motivations, self-confidence, socioeconomic status, and professions of the participants. Indeed, differences in FA between ex-illiterates and illiterates might be present before ex-illiterates voluntarily begin to learn reading. Future studies should attempt to verify whether or not "learning to read improves the structure" or not. To examine the causal changes in the brain, longitudinal studies that track neural indices and behavioral measures, together with multiple regression analyses that de-correlate confounding variables, would be critical (Galván, 2010; Dehaene et al., 2015). Note that our present study does not intend to suggest causal influences of L2 acquisition on structural properties. Rather, we demonstrated here that the performance-related group differences, which were separated from the duration of L2 experience (i.e., DOE) as well as from the age-related group differences, were predicted by FA in the left dorsal pathway. Future studies on adult language acquisition are needed to elucidate hidden but critical issues regarding how non-cognitive factors that drive individual experiences influence the function and structures of the specific brain regions.

Chapter 4. General discussion

When examining individual differences in white matter pathways, it is of great importance to identify each pathway of interest precisely, and then to select regions appropriate for quantification. In this dissertation, I have endeavored to develop semi-automatic methods of defining seeds and selecting ROIs, which were custom-designed to the dorsal and ventral language-related pathways. Using these methods for the Arcuate and IFOF in each hemisphere, I demonstrated that only FA of the left Arcuate was significantly correlated with individual accuracy of the syntactic task. Previous studies using the same Syn and Spe tasks in L2 have shown that the activations in the dorsal F3t of the left IFG, as well as the leftward lateralization of this region's volume, were correlated with the accuracy of the Syn task (Nauchi and Sakai, 2009; Sakai et al., 2009). It has also been shown that the left IFG is specialized in syntactic processing in both L1 and L2 (Sakai et al., 2004; Tatsuno and Sakai, 2005; Sakai et al., 2009). In the Syn-Spe contrast, the left middle temporal gyrus and angular gyrus were also significantly activated (Sakai et al., 2009). In the present study, I further showed that the structural property of the left Arcuate connecting the IFG and parietotemporal regions was correlated with the accuracy of the Syn task.

Further, recruiting monozygotic twins, I examined to what extent their L2 abilities and their structural properties were similar. The present results suggest that the thickness of the left Arcuate is more associated with the shared genetic/environmental factors, whereas both of mutually correlated FA in the left Arcuate and individual syntactic abilities in L2 may be less prone to these shared factors. A recent diffusion MRI study has reported that FA varies along a tract (Johnson et al., 2014). Another study dissected the Arcuate into three segments (long, anterior, and posterior segments), and reported that genetic effects on FA were high for the left long, right anterior, and left posterior segments (Budisavljevic et al., 2015). Here I employed semi-automatic analyses for the diffusion MRI data, thereby examining FA in consistent portions of the tracts, which were not possible in previous studies using the mean FA averaged in the entire tracts. Although the sample size of twin pairs was too small to draw

a stronger conclusion, elaborate studies like the present study would indicate future directions for large-scale studies. Large-scale projects with hundreds of twins, such as Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) and Human Connectome Project (HCP), have reported that FA of major tracts was highly heritable (Sprooten et al., 2014; Kochunov et al., 2015). The present results did not show strong correlations of FA among the twin pairs, but this discrepancy may be due to small sample size. Further studies recruiting both monozygotic and dizygotic twins with better tractography analyses would clarify the contribution of genetic and environmental factors to the language-related pathways. Moreover, methodological improvements on thresholding with less-manual adjustments in seed selection are expected to increase the reliability of diffusion MRI analyses. In the present study, I suggest that some task-related skills and word-level knowledge are largely associated with shared factors between twins, whereas individual syntactic abilities in L2 are less prone to shared factors. I also suggest that the thickness of the left Arcuate may be more associated with shared factors than FA was.

By dividing the participants into subgroups, so that either L2 performance [Junior (High)/Senior (Low)] or age [Senior (Low)/Senior (High)] was matched in group comparisons, I demonstrated differential performance-related and age-related signatures on the left Arcuate. The present results suggest that the left dorsal pathway continued to develop to adolescence, and that performance differences in a syntactic task can be predicted by its FA, independent of age and the duration of experience. A more detailed picture of the developmental mechanism of language-related structures would be provided by closely linking the development of linguistic abilities and that of white matter pathways, as well as the structures of connected cortical regions, which can be examined by such methods as VBM and the myelin mapping technique (Glasser and Van Essen, 2011). Our present study showed both age-related and performance-related group differences in adolescent participants, suggesting the importance of future longitudinal studies with various structural

and functional measurements. In order to combine results obtained by multiple imaging methods, it is necessary to identify cortical regions/structures of interest precisely in the individual brains. In this dissertation, I suggested semi-automatic methods of identifying and analyzing white matter pathways with large individual differences, showing that the structural properties of language-related pathways reflected individual variances in fundamental linguistic abilities, i.e., syntactic abilities. Future studies that employ objective methods of evaluating individual developments would further lead to a selection of tailor-made educational programs.

Conclusions

In this dissertation, I developed semi-automatic methods of defining seeds and selecting ROIs for diffusion MRI analyses, aiming at the functional anatomy of the dorsal and ventral language-related pathways. I performed partial correlation analyses between FA of these pathways and the accuracy of the syntactic task (Experiment I), so that the effects of the accuracy of a spelling task, gender, and handedness, were removed (Chapter 2). Among the two pathways in each hemisphere, only FA of the left Arcuate was significantly correlated with individual accuracy of the syntactic task. Further, I recruited monozygotic twins and examined to what extent their L2 abilities and their structural properties were similar (Experiment II). Within twin pairs, the highest significant correlation was observed for reaction times of the spelling task, while the correlation for the accuracy of the syntactic task was marginal; these two correlation coefficients were significantly different. Moreover, the thickness of the left Arcuate was highly correlated within pairs, while its FA, as well as the thickness/FA in the ventral pathways, was not significantly correlated. The correlation coefficient for the thickness of the left Arcuate was significantly larger than that of the left IFOF. These results suggest that the thickness of the left Arcuate is more associated with the shared genetic/environmental factors, whereas both of mutually correlated FA in the left Arcuate and individual syntactic abilities in L2 may be less prone to these shared factors.

To reveal the differential signatures of performances and age on the plasticity of structural properties in major language-related pathways, I examined individual differences in L2 performances, controlling the duration of experience (Chapter 3). I recruited Japanese students at two ages, i.e., junior (age: 13–14) and senior (age: 16–17) high-school students, all of whom started to expose to English at age 12 or 13. I divided them into subgroups, so that either L2 performance [Junior (High)/Senior (Low)] or age [Senior (Low)/Senior (High)] was matched in group comparisons; the duration of L2 experience was also controlled between the Senior (Low) and Senior (High) groups. I then examined the thickness and FA of

the dorsal and ventral pathways, using semi-automatic methods for selecting regions without branches. Regarding FA in the left Arcuate, the Senior (High) group showed significantly higher FA than the other two groups, indicating performance-related group differences. Moreover, FA in the left Arcuate was selectively correlated with the accuracy of a syntactic task. Regarding the thickness of the left Arcuate, the Senior (High) and Senior (Low) groups showed significantly larger thickness than the Junior (High) group, indicating age-related group differences. These differential performance-related and age-related signatures were evident on the left Arcuate alone, in contrast to the right Arcuate, which showed only mild differences in thickness, and to the bilateral IFOF, which lacked either signature. The present results suggest that the left dorsal pathway continued to develop through adolescence, and that performance differences in a syntactic task were related to individual differences in its FA, independent of age and the duration of experience.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor, Kuniyoshi L. Sakai, for his supervision. His passion for science, as well as advice based on academic experience, always encouraged and inspired me. I would also like to express my appreciation to the members of Sakai Laboratory. I am truly grateful to Hiroyuki Miyashita, Shinri Ohta, Kazuki Iijima, Tomoo Inubushi, Tomoya Nakai, and Ryuta Kinno, for their guidance and support. I also thank Yuichiro Shimizu, Ayaka Tsuchiya, Hayate Tada, Kyouhei Tanaka, Atora Yamada, Tatsuro Kuwamoto, and Keita Umejima for their helpful comments. I wish to express my sincere appreciation to Naoko Komoro for her technical assistance, and to Hiromi Matsuda for her administrative assistance and kind encouragement.

References

Anwander, A., Tittgemeyer, M., Von Cramon, D. Y., Friederici, A. D., and Knösche, T. R. (2007). Connectivity-based parcellation of Broca's area. *Cereb. Cortex* 17, 816-825. doi: 10.1093/cercor/bhk034

Bach, M., Laun, F. B., Leemans, A., Tax, C. M. W., Biessels, G. J., Stieltjes, B., and Maier-Hein, K. H. (2014). Methodological considerations on tract-based spatial statistics (TBSS). *Neuroimage* 100, 358-369. doi: 10.1016/j.neuroimage.2014.06.021

Bornkessel-Schlesewsky, I., Schlesewsky, M., Small, S. L., and Rauschecker, J. P. (2015). Neurobiological roots of language in primate audition: Common computational properties. *Trends Cogn. Sci.* 19, 142-150. doi: 10.1016/j.tics.2014.12.008

Brauer, J., Anwander, A., Perani, D., and Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain Lang.* 127, 289-295. doi: 10.1016/j.bandl.2013.03.001

Brouwer, R. M., Mandl, R. C. W., Peper, J. S., Van Baal, G. C. M., Kahn, R. S., Boomsma, D. I., and Pol, H. E. H. (2010). Heritability of DTI and MTR in nine-year-old children. *Neuroimage* 53, 1085-1092. doi: 10.1016/j.neuroimage.2010.03.017

Brouwer, R. M., Mandl, R. C. W., Schnack, H. G., Van Soelen, I. L. C., Van Baal, G. C., Peper, J. S., Kahn, R. S., Boomsma, D. I., and Pol, H. E. H. (2012). White matter development in early puberty: A longitudinal volumetric and diffusion tensor imaging twin study. *PLOS ONE* 7, e32316, 1-10. doi: 10.1371/journal.pone.0032316

Brown, T. T., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., and Schlaggar, B. L. (2005). Developmental changes in human cerebral functional organization for word generation. *Cereb. Cortex* 15, 275-290. doi: 10.1093/cercor/bhh129

Budisavljevic, S., Dell'acqua, F., Rijsdijk, F. V., Kane, F., Picchioni, M., Mcguire, P., Toulopoulou, T., Georgiades, A., Kalidindi, S., Kravariti, E., Murray, R. M., Murphy, D.

G., Craig, M. C., and Catani, M. (2015). Age-related differences and heritability of the perisylvian language networks. *J. Neurosci.* 35, 12625-12634. doi: 10.1523/jneurosci.1255-14.2015

Catani, M., Allin, M. P. G., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., and Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17163-17168. doi: 10.1073/pnas.0702116104

Catani, M., Dell'acqua, F., Bizzi, A., Forkel, S. J., Williams, S. C., Simmons, A., Murphy, D. G., and Thiebaut De Schotten, M. (2012). Beyond cortical localization in clinico-anatomical correlation. *Cortex* 48, 1262-1287. doi: 10.1016/j.cortex.2012.07.001

Catani, M., Jones, D. K., and Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8-16. doi: 10.1002/ana.20319

Catani, M., and Thiebaut De Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44, 1105-1132. doi: 10.1016/j.cortex.2008.05.004

Chai, X. J., Berken, J. A., Barbeau, E. B., Soles, J., Callahan, M., Chen, J.-K., and Klein, D. (2016). Intrinsic functional connectivity in the adult brain and success in second-language learning. *J. Neurosci.* 36, 755-761. doi: 10.1523/jneurosci.2234-15.2016

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: The MIT Press.

Chung, A. W., Thomas, D. L., Ordidge, R. J., and Clark, C. A. (2013). Diffusion tensor parameters and principal eigenvector coherence: Relation to *b*-value intervals and field strength. *Magn. Reson. Imaging* 31, 742-747. doi: 10.1016/j.mri.2012.11.014

Clark, C. A., and Le Bihan, D. (2000). Water diffusion compartmentation and anisotropy at high b values in the human brain. *Magn. Reson. Med.* 44, 852-859. doi: 10.1002/1522-2594(200012)44:6<852::AID-MRM5>3.0.CO;2-A

Dale, P. S., Harlaar, N., and Plomin, R. (2012). Nature and nurture in school-based second language achievement. *Lang. Learn.* 62, 28-48. doi: 10.1111/j.1467-9922.2012.00705.x

De Champfleur, N. M., Maldonado, I. L., Moritz-Gasser, S., Machi, P., Le Bars, E., Bonafé, A., and Duffau, H. (2013). Middle longitudinal fasciculus delineation within language pathways: A diffusion tensor imaging study in human. *Eur. J. Radiol.* 82, 151-157. doi: 10.1016/j.ejrad.2012.05.034

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234-244. doi: 10.1038/nrn3924

Diedenhofen, B., and Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLOS ONE* 10, e0121945, 1-12. doi: 10.1371/journal.pone.0121945

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Lang. Learn.* 53, 3-32. doi: 10.1111/1467-9922.53222

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Jr., Redfern, B. B., and Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92, 145-177. doi: 10.1016/j.cognition.2003.11.002

Dubois, J., Poupon, C., Thirion, B., Simonnet, H., Kulikova, S., Leroy, F., Hertz-Pannier, L., and Dehaene-Lambertz, G. (2016). Exploring the early organization and maturation of linguistic pathways in the human infant brain. *Cereb. Cortex* 26, 2283-2298. doi: 10.1093/cercor/bhv082

Forkel, S. J., Thiebaut De Schotten, M., Kawadler, J. M., Dell'acqua, F., Danek, A., and Catani, M. (2014). The anatomy of fronto-occipital connections from early blunt dissections to contemporary tractography. *Cortex* 56, 73-84. doi: 10.1016/j.cortex.2012.09.005

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiol. Rev.* 91, 1357-1392. doi: 10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends Cogn. Sci.* 16, 262-268. doi: 10.1016/j.tics.2012.04.001

Fuhrmann, D., Knoll, L. J., and Blakemore, S.-J. (2015). Adolescence as a sensitive period of brain development. *Trends Cogn. Sci.* 19, 558-566. doi: 10.1016/j.tics.2015.07.008

Galantucci, S., Tartaglia, M. C., Wilson, S. M., Henry, M. L., Filippi, M., Agosta, F., Dronkers, N. F., Henry, R. G., Ogar, J. M., Miller, B. L., and Gorno-Tempini, M. L. (2011). White matter damage in primary progressive aphasias: A diffusion tensor tractography study. *Brain* 134, 3011-3029. doi: 10.1093/brain/awr099

Galván, A. (2010). Neural plasticity of development and learning. *Hum. Brain Mapp.* 31, 879-890. doi: 10.1002/hbm.21029

Gardner, R. C. (2010). *Motivation and Second Language Acquisition: The Socio-Educational Model.* New York, NY: Peter Lang.

Glasser, M. F., and Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cereb. Cortex* 18, 2471-2482. doi: 10.1093/cercor/bhn011

Glasser, M. F., and Van Essen, D. C. (2011). Mapping human cortical areas *in vivo* based on myelin content as revealed by T1- and T2-weighted MRI. *J. Neurosci.* 31, 11597-11616. doi: 10.1523/jneurosci.2180-11.2011

Griffiths, J. D., Marslen-Wilson, W. D., Stamatakis, E. A., and Tyler, L. K. (2013). Functional organization of the neural language system: Dorsal and ventral pathways are critical for syntax. *Cereb. Cortex* 23, 139-147. doi: 10.1093/cercor/bhr386

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends Cogn. Sci.* 9, 416-423. doi: 10.1016/j.tics.2005.07.004

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Curr. Opin. Neurobiol.* 28, 136-141. doi: 10.1016/j.conb.2014.07.013

Hickok, G., and Poeppel, D. (2007). Opinion - The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393-402. doi: 10.1038/nrn2113

Homae, F., Hashimoto, R., Nakajima, K., Miyashita, Y., and Sakai, K. L. (2002). From perception to sentence comprehension: The convergence of auditory and visual information of language in the left inferior frontal cortex. *Neuroimage* 16, 883-900. doi: 10.1006/nimg.2002.1138

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., and Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: A multimodal imaging study. *J. Neurosci.* 33, 13663-13672. doi: 10.1523/jneurosci.0410-13.2013

Iijima, K., Fukui, N., and Sakai, K. L. (2009). The cortical dynamics in building syntactic structures of sentences: An MEG study in a minimal-pair paradigm. *Neuroimage* 44, 1387-1396. doi: 10.1016/j.neuroimage_2008.10.041

Iijima, K., and Sakai, K. L. (2014). Subliminal enhancement of predictive effects during syntactic processing in the left inferior frontal gyrus: An MEG study. *Front. Syst. Neurosci.* 8, 217, 1-14. doi: 10.3389/fnsys.2014.00217

Im, K., Pienaar, R., Lee, J.-M., Seong, J.-K., Choi, Y. Y., Lee, K. H., and Grant, P. E. (2011). Quantitative comparison and analysis of sulcal patterns using sulcal graph matching: A twin study. *Neuroimage* 57, 1077-1086. doi: 10.1016/j.neuroimage.2011.04.062

Inubushi, T., Iijima, K., Koizumi, M., and Sakai, K. L. (2012). Left inferior frontal activations depending on the canonicity determined by the argument structures of ditransitive sentences: An MEG study. *PLOS ONE* 7, e37192, 1-11. doi: 10.1731/journal.pone.0037192

Johnson, R. T., Yeatman, J. D., Wandell, B. A., Buonocore, M. H., Amaral, D. G., and Nordahl, C. W. (2014). Diffusion properties of major white matter tracts in young, typically developing children. *Neuroimage* 88, 143-154. doi: 10.1016/j.neuroimage.2013.11.025

Jones, D. K., and Basser, P. J. (2004). “Squashing peanuts and smashing pumpkins”: How

noise distorts diffusion-weighted MR data. *Magn. Reson. Med.* 52, 979-993. doi: 10.1002/mrm.20283

Kinno, R., Ohta, S., Muragaki, Y., Maruyama, T., and Sakai, K. L. (2014). Differential reorganization of three syntax-related networks induced by a left frontal glioma. *Brain* 137, 1193-1212. doi: 10.1093/brain/awu013

Kochunov, P., Jahanshad, N., Marcus, D., Winkler, A., Sprooten, E., Nichols, T. E., Wright, S. N., Hong, L. E., Patel, B., Behrens, T., Jbabdi, S., Andersson, J., Lenglet, C., Yacoub, E., Moeller, S., Auerbach, E., Ugurbil, K., Sotiropoulos, S. N., Brouwer, R. M., Landman, B., Lemaitre, H., Den Braber, A., Zwiers, M. P., Ritchie, S., Van Hulzen, K., Almasy, L., Curran, J., Dezubicaray, G. I., Duggirala, R., Fox, P., Martin, N. G., Mcmahon, K. L., Mitchell, B., Olvera, R. L., Peterson, C., Starr, J., Sussmann, J., Wardlaw, J., Wright, M., Boomsma, D. I., Kahn, R., De Geus, E. J. C., Williamson, D. E., Hariri, A., Van 'T Ent, D., Bastin, M. E., Mcintosh, A., Deary, I. J., Hulshoff Pol, H. E., Blangero, J., Thompson, P. M., Glahn, D. C., and Van Essen, D. C. (2015). Heritability of fractional anisotropy in human white matter: A comparison of Human Connectome Project and ENIGMA-DTI data. *Neuroimage* 111, 300-311. doi: 10.1016/j.neuroimage.2015.02.050

Kuhl, P. K., Tsao, F.-M., and Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9096-9101. doi: 10.1073/pnas.1532872100

Lebel, C., and Beaulieu, C. (2009). Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Hum. Brain Mapp.* 30, 3563-3573. doi: 10.1002/hbm.20779

Lebel, C., and Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *J. Neurosci.* 31, 10937-10947. doi: 10.1523/jneurosci.5302-10.2011

Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second

language learning: Anatomical changes in the human brain. *Cortex* 58, 301-324. doi: 10.1016/j.cortex.2014.05.001

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., and Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Lang. Learn.* 63, 530-566. doi: 10.1111/lang.12011

López-Barroso, D., Catani, M., Ripollés, P., Dell'acqua, F., Rodríguez-Fornells, A., and De Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13168-13173. doi: 10.1073/pnas.1301696110

Magnusdottir, S., Fillmore, P., Den Ouden, D. B., Hjaltason, H., Rorden, C., Kjartansson, O., Bonilha, L., and Fridriksson, J. (2013). Damage to left anterior temporal cortex predicts impairment of complex syntactic processing: A lesion-symptom mapping study. *Hum. Brain Mapp.* 34, 2715-2723. doi: 10.1002/hbm.22096

Makris, N., and Pandya, D. N. (2009). The extreme capsule in humans and rethinking of the language circuitry. *Brain Struct. Funct.* 213, 343-358. doi: 10.1007/s00429-008-0199-8

Martino, J., Brogna, C., Robles, S. G., Vergani, F., and Duffau, H. (2010). Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. *Cortex* 46, 691-699. doi: 10.1016/j.cortex.2009.07.015

Morris, D. M., Embleton, K. V., and Parker, G. J. M. (2008). Probabilistic fibre tracking: Differentiation of connections from chance events. *Neuroimage* 42, 1329-1339. doi: 10.1016/j.neuroimage.2008.06.012

Nauchi, A., and Sakai, K. L. (2009). Greater leftward lateralization of the inferior frontal gyrus in second language learners with higher syntactic abilities. *Hum. Brain Mapp.* 30, 3625-3635. doi: 10.1002/hbm.20790

Ohta, S., Fukui, N., and Sakai, K. L. (2013). Syntactic computation in the human brain: The Degree of Merger as a key factor. *PLOS ONE* 8, e56230, 1-16. doi: 10.1371/journal.pone.0056230

Oishi, K., Faria, A., Van Zijl, P. C. M., and Mori, S. (2011). *MRI Atlas of Human White Matter.* London: Academic Press.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9, 97-113. doi: 10.1016/0028-3932(71)90067-4

Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., and Friederici, A. D. (2011). Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16056-16061. doi: 10.1073/pnas.1102991108

Raettig, T., Frisch, S., Friederici, A. D., and Kotz, S. A. (2010). Neural correlates of morphosyntactic and verb-argument structure processing: An EfMRI study. *Cortex* 46, 613-620. doi: 10.1016/j.cortex.2009.06.003

Sakai, K. L., Miura, K., Narafu, N., and Muraishi, M. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cereb. Cortex* 14, 1233-1239. doi: 10.1093/cercor/bhh084

Sakai, K. L., Nauchi, A., Tatsuno, Y., Hirano, K., Muraishi, M., Kimura, M., Bostwick, M., and Yusa, N. (2009). Distinct roles of left inferior frontal regions that explain individual differences in second language acquisition. *Hum. Brain Mapp.* 30, 2440-2452. doi: 10.1002/hbm.20681

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., and Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18035-18040. doi: 10.1073/pnas.0805234105

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D. E., and Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nat. Neurosci.* 19, 1250-1255. doi: 10.1038/nn.4354

Schlaggar, B. L., Brown, T. T., Lugar, H. M., Visscher, K. M., Miezin, F. M., and Petersen, S.

E. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science* 296, 1476-1479. doi:

Schlegel, A. A., Rudelson, J. J., and Tse, P. U. (2012). White matter structure changes as adults learn a second language. *J. Cogn. Neurosci.* 24, 1664-1670. doi: 10.1162/jocn_a_00240

Scholz, J., Klein, M. C., Behrens, T. E. J., and Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nat. Neurosci.* 12, 1370-1371. doi: 10.1038/nn.2412

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591-611. doi: 10.2307/2333709

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., and Behrens, T. E. J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31, 1487-1505. doi: 10.1016/j.neuroimage.2006.02.024

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., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208-S219. doi: 10.1016/j.neuroimage.2004.07.051

Sotiropoulos, S. N., Moeller, S., Jbabdi, S., Xu, J., Andersson, J. L., Auerbach, E. J., Yacoub, E., Feinberg, D., Setsompop, K., Wald, L. L., Behrens, T. E. J., Ugurbil, K., and Lenglet, C. (2013). Effects of image reconstruction on fiber orientation mapping from multichannel diffusion MRI: Reducing the noise floor using SENSE. *Magn. Reson. Med.* 70, 1682-1689. doi: 10.1002/mrm.24623

Sprooten, E., Knowles, E. E., Mckay, D. R., Göring, H. H., Curran, J. E., Kent Jr, J. W., Carless, M. A., Dyer, T. D., Drigalenko, E. I., Olvera, R. L., Fox, P. T., Almasy, L., Duggirala, R., Kochunov, P., Blangero, J., and Glahn, D. C. (2014). Common genetic

variants and gene expression associated with white matter microstructure in the human brain. *Neuroimage* 97, 252-261. doi: 10.1016/j.neuroimage.2014.04.021

Suzuki, K., and Sakai, K. L. (2003). An event-related fMRI study of explicit syntactic processing of normal/anomalous sentences in contrast to implicit syntactic processing. *Cereb. Cortex* 13, 517-526. doi: 10.1093/cercor/13.5.517

Tatsuno, Y., and Sakai, K. L. (2005). Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *J. Neurosci.* 25, 1637-1644. doi: 10.1523/jneurosci.3978-04.2005

Thiebaut De Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., and Dehaene, S. (2014). Learning to read improves the structure of the arcuate fasciculus. *Cereb. Cortex* 24, 989-995. doi: 10.1093/cercor/bhs383

Thiebaut De Schotten, M., Dell'acqua, F., Valabregue, R., and Catani, M. (2012). Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* 48, 82-96. doi: 10.1016/j.cortex.2011.10.001

Thiebaut De Schotten, M., Ffytche, D. H., Bizzi, A., Dell'acqua, F., Allin, M., Walshe, M., Murray, R., Williams, S. C., Murphy, D. G. M., and Catani, M. (2011). Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. *Neuroimage* 54, 49-59. doi: 10.1016/j.neuroimage.2010.07.055

Thompson, P. M., Cannon, T. D., Narr, K. L., Van Erp, T., Poutanen, V.-P., Huttunen, M., Lönnqvist, J., Standertskjöld-Nordenstam, C.-G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C. I., and Toga, A. W. (2001). Genetic influences on brain structure. *Nat. Neurosci.* 4, 1253-1258. doi: 10.1038/nn758

Thothathiri, M., Kimberg, D. Y., and Schwartz, M. F. (2012). The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion symptom mapping in aphasia. *J. Cogn. Neurosci.* 24, 212-222. doi: 10.1162/jocn_a_00118

Tsang, J. M., Dougherty, R. F., Deutsch, G. K., Wandell, B. A., and Ben-Shachar, M. (2009). Frontoparietal white matter diffusion properties predict mental arithmetic skills in children. *Proc. Natl. Acad. Sci. U.S.A.* 106, 22546-22551. doi: 10.1073/pnas.0906094106

Tuch, D. S. (2004). Q-ball imaging. *Magn. Reson. Med.* 52, 1358-1372. doi: 10.1002/mrm.20279

Van Hecke, W., Leemans, A., De Backer, S., Jeurissen, B., Parizel, P. M., and Sijbers, J. (2010). Comparing isotropic and anisotropic smoothing for voxel-based DTI analyses: A simulation study. *Hum. Brain Mapp.* 31, 98-114. doi: 10.1002/hbm.20848

White, T., Andreasen, N. C., and Nopoulos, P. (2002). Brain volumes and surface morphology in monozygotic twins. *Cereb. Cortex* 12, 486-493. doi: 10.1093/cercor/12.5.486

Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., Ogar, J. M., Deleon, J., Miller, B. L., and Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron* 72, 397-403. doi: 10.1016/j.neuron.2011.09.014

Yamamoto, K., and Sakai, K. L. (2016). The dorsal rather than ventral pathway better reflects individual syntactic abilities in second language. *Front. Hum. Neurosci.* 10, 295, 1-18. doi: 10.3389/fnhum.2016.00295

Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., and Feldman, H. M. (2012). Tract Profiles of white matter properties: Automating fiber-tract quantification. *PLOS ONE* 7, e49790, 1-15. doi: 10.1371/journal.pone.0049790

Yeatman, J. D., Dougherty, R. F., Rykhlevskaia, E., Sherbondy, A. J., Deutsch, G. K., Wandell, B. A., and Ben-Shachar, M. (2011). Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. *J. Cogn. Neurosci.* 23, 3304-3317. doi: 10.1162/jocn_a_00061

Yusa, N. (2003). "'Passive' unaccusatives in L2 acquisition", in *Japanese/Korean Linguistics, vol. 11*, ed. P.M. Clancy (Stanford, CA: Center for the Study of Language and Information

Publications and Stanford Linguistic Association), 246-259.