Behavioral/Cognitive

Dynamic Neural Network Reorganization Associated with Second Language Vocabulary Acquisition: A Multimodal Imaging Study

Chihiro Hosoda,[1,2,3,4] **Kanji Tanaka,**[5,6] **Tadashi Nariai,**[3] **Manabu Honda,**[1,2] and **Takashi Hanakawa**[1,2,7]

[1]Department of Functional Brain Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo 187-8502, Japan, [2]Department of Advanced Neuroimaging, Integrative Brain Imaging Center, National Center of Neurology and Psychiatry, Tokyo 187-8551, Japan, [3]Department of Neurosurgery, Tokyo Medical and Dental University, Tokyo 113-0034, Japan, [4]Department of Motor Control and Rehabilitation, ATR Computational Neuroscience Laboratories, Kyoto, 619-0288, Japan, [5]Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, 113-8654, Japan [6]Japan Society for the Promotion of Science, Tokyo, 102-0083, Japan, and [7]PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan

It remains unsettled whether human language relies exclusively on innately privileged brain structure in the left hemisphere or is more flexibly shaped through experiences, which induce neuroplastic changes in potentially relevant neural circuits. Here we show that learning of second language (L2) vocabulary and its cessation can induce bidirectional changes in the mirror-reverse of the traditional language areas. A cross-sectional study identified that gray matter volume in the inferior frontal gyrus pars opercularis (IFGop) and connectivity of the IFGop with the caudate nucleus and the superior temporal gyrus/supramarginal (STG/SMG), predominantly in the right hemisphere, were positively correlated with L2 vocabulary competence. We then implemented a cohort study involving 16 weeks of L2 training in university students. Brain structure before training did not predict the later gain in L2 ability. However, training intervention did increase IFGop volume and reorganization of white matter including the IFGop-caudate and IFGop-STG/SMG pathways in the right hemisphere. These "positive" plastic changes were correlated with the gain in L2 ability in the trained group but were not observed in the control group. We propose that the right hemispheric network can be reorganized into language-related areas through use-dependent plasticity in young adults, reflecting a repertoire of flexible reorganization of the neural substrates responding to linguistic experiences.

Introduction

Language relies on both innate ability and experiences (Kuhl, 2010; Hsu et al., 2011), making understanding of experience-dependent shaping of language systems an essential theme for language neuroscience. The neural mechanisms of experienced-induced first language (L1) development are difficult to study since L1 is acquired through infancy-childhood. Understanding of second language (L2) learning is therefore important because it should also invoke basic mechanisms of language acquisition.

Vocabulary is important for proficiency in any language (Coady and Huckin, 1997; Grogan et al., 2012), but neural mecha-

nisms of L2 vocabulary learning remain unclear. L2 vocabulary learning is achieved through interactions with the already existing L1 lexicon in adults (Ellis et al., 1999). For example, L2 stimuli spontaneously activate the L1 lexicon (Thierry and Wu, 2007), and individuals with high L1 vocabulary competence have an advantage in L2 vocabulary learning (Meschyan, 2002). However, despite these interactions, the L1 and L2 lexicons may be differentially structured; it has been proposed that the L2 lexicon is more phonologically organized compared with the L1 lexicon strongly connected with the semantic system (Laufer, 1989). Hence, L2 vocabulary leaning in adults would induce reorganization predominantly in the phonological system nested with the semantic system. These features may make L2 vocabulary learning in adults a unique paradigm for exploring repertoires of dynamic neural reorganization of the language system, especially in the case of linguistically distant languages such as English and Japanese (Chiswick and Miller, 2005).

L2 learning may accompany macroscopic structural alterations of the language system. Recent neuroimaging studies have revealed plastic changes of gray matter (GM) during learning of various cognitive and motor abilities (Draganski et al., 2004; Mechelli et al., 2004). The analysis of fractional anisotropy (FA) computed from diffusion-weighted magnetic resonance imaging

Table 1. Means and SDs of subject demographics, L2 proficiency tests, IQ, and personality assessments (NEO-FFI)

	Cross-sectional study (n = 137)	Cohort study (pretraining data)	
		TG (n = 24)	CG (n = 20)
Age	21.8 ± 16.0	20.1 ± 1.8	20.5 ± 1.7
Sex (male: female)	66:71	10:14	10:10
TOEIC (200)	NA	102.5 ± 2.6	100.8 ± 2.3
EVT (100)	20.6 ± 16.0	11.3 ± 10.9	12.9 ± 12.1
NART (100)	41.8 ± 6.2	35.6 ± 4.5	39.8 ± 4.0
WAIS-TIQ	NA	113.0 ± 2.2	116.5 ± 1.8
WAIS-VIQ	NA	101.0 ± 2.6	104.0 ± 2.1
WAIS-PIQ	NA	103.5 ± 2.5	101.0 ± 2.3
NEO-FFI neuroticism	NA	25.4 ± 1.4	26.4 ± 1.4
NEO-FFI extraversion	NA	29.8 ± 1.4	30.6 ± 1.3
NEO-FFI openness	NA	32.2 ± 1.3	32.3 ± 1.3
NEO-FFI agreeableness	NA	33.9 ± 1.0	34.1 ± 1.1
NEO-FFI conscientiousness	NA	31.5 ± 1.3	30.7 ± 1.2

In the cohort study, no significant differences were found between the TG and CG for all measurement types. NA, not available; WAIS-VIQ, Wechsler Adult Intelligence Scale Verbal IQ; WAIS-TIQ, Wechsler Adult Intelligence Scale Total IQ; WAIS-PIQ, Wechsler Adult Intelligence Scale Performance IQ; NEO-FFI, NEO Five-Factor Inventory.

(DWI) can index organization of white matter (WM) microstructure, and allows us to visualize plastic changes of WM (Johansen-Berg, 2012; Zatorre et al., 2012). Very recently, a few L2 learning studies have shown GM (Mårtensson et al., 2012) and WM (Schlegel et al., 2012) changes in the traditional language network in the left hemisphere.

Previous neuroimaging studies have, however, demonstrated roles of the right hemisphere in language ability (Vingerhoets et al., 2003; Videsott et al., 2010; Vigneau et al., 2011; Van Ettinger-Veenstra et al., 2012) and in L2 control (Hernandez et al., 2001; Hosoda et al., 2012). One of these found correlations between right prefrontal activity and L2 vocabulary competence (Hosoda et al., 2012). To explore neural plasticity associated with L2 vocabulary learning, we first explored the neural underpinnings correlated with L2 vocabulary competence in a cross-sectional study, revealing GM and WM structures correlated with L2 vocabulary levels predominantly in the right hemisphere. We then ran a cohort study implementing an L2 training program to test whether those L2 vocabulary correlates reflected innate predispositions or plastic changes resulting from learning. We further examined the natural course of the induced structural changes. Here we provide evidence that L2 vocabulary learning induces dynamic reorganization of GM and WM structures outside of the typical language network in adult brains.

Materials and Methods

Subjects. In the cross-sectional study, 137 native Japanese speakers (71 males and 66 females) with a mean age of 24.0 years [SD = 5.3, range 18–42] were enrolled (Table 1). All subjects were university students or graduates whose self-reported English (L2) proficiencies varied from low to very high to the level of Japanese–English bilinguals.

In the cohort study, 67 native Japanese speakers (mean age = 20.1, 31 men and 36 women) were initially placed into a training group (TG). In the TG, 24 of 47 participants completed the planned 4 month training, whereas the remaining 23 participants opted not to continue during the course of training for various reasons. Hence, we studied 24 TG participants (mean age = 20.1, 10 men and 14 women) and an age-matched control group (CG; n = 20; mean age = 20.1, 10 men and 10 women). The participants in the cohort study were university students and recruited from basic level English classes (Table 1). The participants in the cross-sectional study and those in the cohort study did not overlap.

In both studies, the participants had grown up in Japan and started to learn English as their L2 at a mean age of 11.0 years. None had started to learn L2 before 7 years old, and thus all participants were regarded as late L2 learners (Dowens et al., 2010). All the participants were right-handed

as assessed by the Edinburgh Handedness Inventory. All of the participants showed handedness scale of 100 (SD = 0) because we only recruited the perfectly right-handed participants, revealing the contents of handedness inventory in the recruitment information. All were healthy and neurologically intact, with no history of neuropsychiatric disorders, psychotropic medication use, or head injury. All the participants gave written informed consent according to the study protocol approved by the institutional review board (National Center of Neurology and Psychiatry).

Experimental design. In the cross-sectional study, the participants underwent MRI scanning (T1-weighted images and DWI) and assessment for English proficiency using the English Vocabulary Test (EVT) and the National Adult Reading Test (NART). The T1-weighted MRI provided the data for voxel-based morphometry (VBM) analysis of the GM volume, and the DWI allowed for tract-based spatial statistics (TBSS) and probabilistic diffusion-based tractography (PDT).

In the cohort experiment, we adopted the Test of English for International Communication (TOEIC) as a primary measure to evaluate various L2 abilities. The TOEIC assesses English proficiency for use in business, providing a fairly accurate measurement of English capabilities for non-native speakers in listening, reading, and grammar. The EVT and NART were used as adjunctive measures of L2 proficiency. Additionally, the Wechsler Adult Intelligence Scale-3 (WAIS-3) and NEO-Five Factor Inventory (NEO-FFI) were used to confirm homogeneity of basic intelligence and personality traits between the TG and CG.

All participants underwent MRI scanning (T1-weighted MRI and DWI) and an L2 assessment battery before training, hereafter referred to as the "Pre" condition. The L2 training program was developed in-house using Visual Basic. Participants undertook 16 weeks of training, and in each week they learned 60 words or idioms including the meaning, spelling, and pronunciation as well as example sentences indicating their usage. For pronunciation, we referred to the AT&T Natural Voices Text-to-Speech Demo (http://www2.research.att.com/~ttsweb/tts/demo.php). The participants were encouraged to dictate each word, idiom, and sentence 10 times. Each weekend, we distributed a program that included a review test using the 60 words or idioms of the week. Upon completing the full program, the participants were expected to master almost 1000 words and idioms. Participants in the CG received no particular assignment according to previous cohort studies using e-learning (Takeuchi et al., 2010; Mårtensson et al., 2012; Schlegel et al., 2012; Wan et al., 2012; Ghazi Saidi et al., 2013).

The participants in both the TG and CG underwent an L2 assessment battery immediately after the training period ("Post-1"). After completion of the training program, the participants were not further engaged in specific L2 learning programs systemically, and thus the continuation of L2 learning depended upon each participant's choice in personal life. We obtained follow-up behavioral and imaging data from the TG group a year after the end of the training program ("Post-2"). Statistical tests on the behavioral parameters were performed using SPSS 17.0 (IBM). To explore training-induced changes in L2 abilities, we analyzed total TOEIC score and its subsections (listening, reading, and grammar) using a 2-by-2 mixed repeated-measures ANOVA with the time (Pre and Post-1) as a within-subject variable and the group (TG and CG) as a between-subject variable. Next, a correlation analysis was performed to test if the training-related changes in TOEIC score were correlated with those of EVT or NART score.

Image data acquisition. MRI data were acquired using a 3 T MRI scanner (Siemens Trio) with an 8-channel phased array receiver coil. High-resolution, 3D, T1-weighted anatomical images were obtained with an MPRAGE sequence designed as follows: repetition time (TR) = 2000 ms, echo time (TE) = 4.4 ms, inversion time (TI) = 990 ms, flip angle = 80°, matrix size = 192 × 176, field of view (FOV) = 192 × 176 mm and 1 mm³ isotropic voxels. We also acquired whole-brain DWI as follows: TR = 7900 ms, TE = 80 ms, 65 slices, flip angle = 90°, matrix size = 96 × 96, FOV = 192 × 192 mm, 2 × 2 × 2 mm³ isotropic voxels, 81 volumes with diffusion weighting (b value = 700 s/mm²) for different motion probing gradient directions and 9 volumes without diffusion weighting (b = 0 s/mm²). Field-map images were acquired in the same scanning space as the DWI (TE1 = 5.19 ms; TE2 = 7.65 ms). All image data were

converted into the Neuroimaging Informatics Technology Initiative format before further processing.

Image data analysis: GM-VBM. The high-resolution 3D T1-weighted images were subjected to a VBM analysis using VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm.html) implemented in SPM8 (http://www.fil.ion.ucl.ac.uk/spm). The preprocessing steps were as follows. (1) Each image was segmented into GM, WM, and CSF images in the native image space. (2) The diffeomorphic anatomical registration through exponentiated Lie algebra (DARTEL) registration method was used to create a study-specific mean GM image template by using the aligned images from all the subjects to improve intersubject registration (Ashburner, 2007). Individual's GM images were registered to the study-specific mean GM template. (4) The registered GM images were further transformed into the Montreal Neurological Institute (MNI) space. (5) These normalized GM images were smoothed using a Gaussian kernel of 12 mm full-width at half-maximum.

In the cross-sectional study, we investigated the correlations of EVT or NART scores with regional GM volume. A multiple regression design was used using EVT or NART score as an independent variable and sex, duration of L2 learning, and age of acquisition as nuisance variables. We set the voxelwise significance level at $p < 0.05$ corrected for multiple comparisons in terms of the familywise error (FWE) rate. After identifying the inferior fontal gyrus pars opercularis (IFGop), superior temporal gyrus (STG), and caudate nucleus as the correlates of L2 vocabulary, we further investigated which hemisphere was more relevant. We computed the laterality index (LI) of GM volume correlated with L2 vocabulary competence (EVT), by using mean β-values computed from 5 mm spherical volumes of interest (VOIs) in IFGop ($x, y, z = \pm 40, 9, 24$), STG/supramarginal gyrus (SMG; $x, y, z = \pm 57, -36, 6$), and caudate nucleus as follows ($x, y, z = \pm 6, 8, 8$). All of these VOIs were defined in the left hemisphere, and thus mirror-reversed VOIs were created for the right hemisphere.

$$LI = (mean\beta_{LEFT\_VOI} - mean\beta_{RIGHT\_VOI})/$$

$$(mean\beta_{LEFT\_VOI} + mean\beta_{RIGHT\_VOI})$$

The selection of VOI (IFGop, STG, and caudate nucleus) was based on the results from previous studies (Crinion et al., 2006; Saur et al., 2008). The LI ranges from -1 (completely lateralized to the right) to $+1$ (completely lateralized to the left). Individuals with an LI of $> +0.4$ or < -0.4 were categorized into the left or right hemisphere dominant group, respectively, while those with LI between -0.4 and $+0.4$ were placed into the bilateral representation group (Briellmann et al., 2003).

In the cohort study, we first tested the possibility that GM volume before L2 training could predict the degree of L2 ability improvement after training. Positive results from this analysis would support the hypothesis that particular brain regions are reserved for L2 vocabulary learning. If such were the case, people with particularly developed structures somewhere in the brain might show better L2 learning ability. To test this hypothesis, a correlation analysis was performed using the MRI data of the Pre condition and the training-related improvements in L2 ability (differences in TOEIC scores between the Pre and Post-1 conditions). Next, to explore training-induced changes in GM, we conducted a 2-by-2 mixed repeated-measures ANOVA with time (Pre and Post-1) as a within-subject variable and group (TG and CG) as a between-subject variable. The regions showing significant time-by-group interaction ($p < 0.05$, FWE corrected) were explored. Third, a correlation analysis was performed to test if the training-related changes in TOEIC score were correlated with those of GM volume in right IFGop. This analysis tested if individual differences in brain plastic changes accounted for individual differences in L2 performance improvement after training.

Image data analysis: DWI-TBSS. Data preprocessing and analysis of DWI were performed using Oxford Centre for Functional MRI of the Brain (FMRIB)'s software library (FSL 4.1, UK; http://www.fmrib.ox.ac.uk/fsl/]. All DWI were registered to the b0 images. Nonlinear image distortions due to magnetic field (b0) inhomogeneity were corrected based on the field map images, using FUGUE in the FMRIB software library. The registered images were skull stripped using the Brain Extraction Tool. FA maps were calculated using the FMRIB's Diffusion Tool-

box (FDT) v2.0 (Smith et al., 2004). After calculation of the FA map for each subject, we implemented a voxelwise statistical analysis of the FA data using TBSS v1.2 (Smith et al., 2006). In brief, TBSS was performed as follows: (1) alignment of the FA images from all subjects to a template that was arbitrarily selected from those FA images, (2) transformation of all the aligned FA images into MNI space using affine registration to remove the effect of cross-subject spatial variability (Kerns et al., 2004), (3) creation of a mean FA image and FA skeleton images corresponding to the center of the WM using a threshold of FA ≥ 0.20 (Smith et al., 2006), (4) projection of the individuals' FAs onto the mean FA skeleton, and (5) voxelwise cross-subject statistical analyses.

As water molecules move faster along the direction of WM fibers than in the direction perpendicular to them, the direction and size of water diffusion in the brain can index structural organization of WM fibers (Le Bihan and Johansen-Berg, 2012; Zatorre et al., 2012). DWI can measure water diffusion noninvasively *in vivo*. FA is a parameter computed from DWI according to tensor-based modeling of water diffusion, and represents the degree of directional bias of water diffusion. FA is sensitive to size and density of axons, degrees of myelination, and the coherence of organization of fibers within a voxel (Beaulieu, 2002; Alexander et al., 2007; Zatorre et al., 2012) In the cross-sectional study, correlations of L2 proficiency (EVT or NART) with FA values as an index of WM organization were tested using Randomize v2.1 in FSL. The statistical threshold was set at $p < 0.05$ (FWE corrected), and the threshold-free cluster enhancement method was used to define the clusters.

In the cohort study, similar to the VBM analysis, we first tested the possibility that structural organization of WM before L2 training could predict the degree of L2 ability improvement after training. Recent literature indicates that training-induced changes of FA can capture plastic reorganization of WM (Johansen-Berg, 2007, 2012; Flöel et al., 2009; Scholz et al., 2009; Johansen-Berg et al., 2012; Zatorre et al., 2012). Activity-dependent myelo-modulation is a potential mechanism by which WM is carved by experiences. Such structural WM changes, including production of myelin basic protein, are correlated with changes in FA (Blumenfeld-Katzir et al., 2011). Changes in myelination may alter conduction velocity and synchronization of signal transmission across the remote areas (Fields, 2005), relating probably to behavioral changes. To capture such learning-induced reorganization of WM, we conducted a 2-by-2 mixed repeated-measures ANOVA with time as a within-subject variable and the group as a between-subject variable ($p < 0.05$, FWE corrected), yielding FA changes specific to the L2 training program. A correlation analysis was performed using FA images (Pre) and the changes in TOEIC scores from the Pre to Post-1 conditions.

Image analysis: multifiber PDT. We conducted PDT using FDT (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT). We performed tractography by tracing pathways through the estimates of diffusion directions. A multifiber model was used to handle the issue of fiber crossing. In each voxel, two fiber directions were modeled, and a probability distribution function (pdf) of diffusion parameters was estimated. We then proceeded with PDT, drawing multiple (5000) streamline samples based on the pdf and estimated the distribution of connections from each seed to target voxels. The generated pathways were volumes in which values at each voxel represented the number of samples passing through that voxel. The voxel value corresponded to the probability of the target voxel connecting to the seed voxel. To remove spurious nonsignificant connections, pathways in individual subjects were thresholded to include only voxels that had at least 50 samples passing through them (out of 5000 samples drawn from each seed voxel).

Previous studies have pointed to the role of the caudate nucleus in L2 (Crinion et al., 2006; Abutalebi et al., 2008; Grogan et al., 2009; Hosoda et al., 2012; Mårtensson et al., 2012). Moreover, the present TBSS result suggested a correlation between L2 ability and WM situated between the IFG and the caudate nucleus and temporoparietal junction (STG/SMG). In the cross-sectional study, therefore, a correlation analysis was performed to investigate the correlation between an L2 proficiency index (EVT or NART score) and the WM connectivity parameter computed from the bilateral IFGop-caudate and IFGop-STG/SMG corresponding to the dorsal pathway (Rilling et al., 2008; Saur et al., 2008). The whole caudate nucleus was manually defined for a target mask in each hemi-
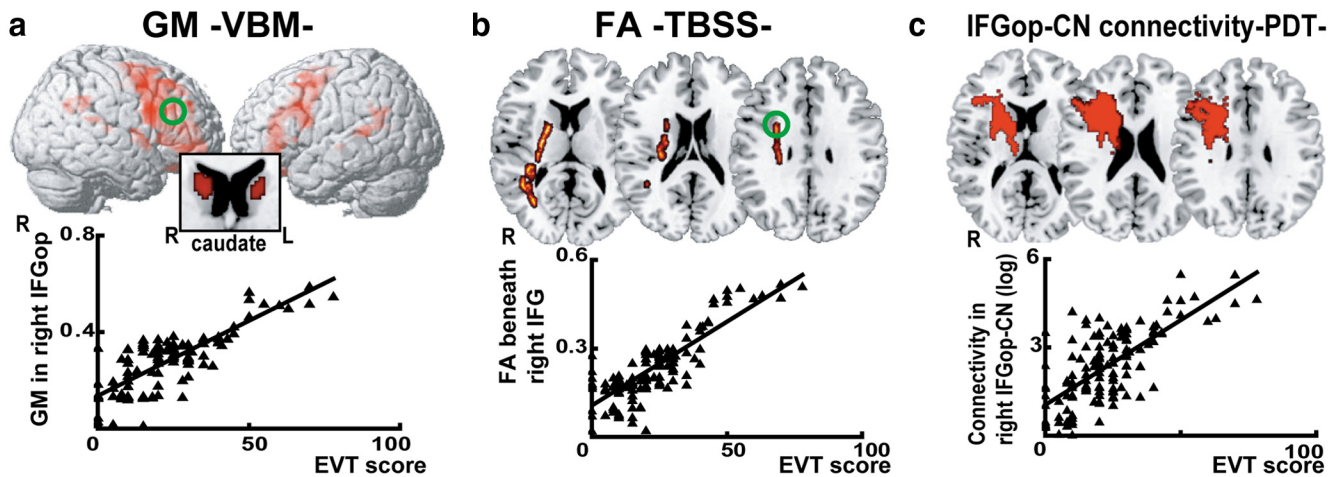
Figure 1. Results from the cross-sectional study. ***a***, 3D rendered images showing the correlation between EVT score and GM volume from the VBM analysis. GM volume for the bilateral IFGop, STG/SMG, and caudate nucleus (CN) correlated with EVT score, reflecting differential levels of L2 proficiency across participants (*p* < 0.05 FWE corrected). The plots show the correlation between EVT and GM volume in the right IFGop (green circle). ***b***, The TBSS analysis displays voxels with a positive correlation between EVT score and fractional anisotropy reflecting integration of WM structure. Significant correlation was found in the subcortical region beneath the IFGop and the AF. The plots show the correlation between EVT scores and FA values in the WM beneath the right IFGop (green circle). ***c***, The PDT analysis shows fiber connections between the right IFGop and the CN in all participants. The EVT score was significantly correlated with connectivity of the IFGop-caudate head (plot). R, right; L, left.

sphere according to each subject's 3D T1-weighted image (Peltier et al., 2011). A spherical VOI with a 5 mm radius was used as a seed mask for the IFGop. The dorsal pathway corresponds to fibers connecting IFGop and STG, and includes arcuate fasciculus (AF) and parts of superior longitudinal fasciculus (SLF; Rilling et al., 2008; Saur et al., 2008). According to a previous study (Saur et al., 2008), we implemented 5 mm radius spherical VOIs in the bilateral IFGop and STG/SMG for the dorsal pathway. We performed PDT within each hemisphere: IFGop-caudate nucleus and IFGop-STG/SMG (dorsal pathway) (four tracts in total).

In the cohort study, in addition to the IFGop-caudate and IFGop-STG/SMG pathways, we also assessed correlations between the L2 competence and other long frontoparietal/temporal association tracts that are reported to be relevant to language processing/learning. We considered the ventral language pathways and inferior longitudinal fasciculus (ILF). The ventral pathway connects IFG pars triangularis (IFGtri) and middle temporal gyrus (MTG), passing through the extreme capsule. ILF connects medial temporal lobe with the temporo-occipital junction, and is suggested be relevant to lexical-semantic language processing (Catani and Mesulam, 2008; Rilling et al., 2008; Saur et al., 2008). We implemented 5 mm radius spherical VOIs in the bilateral IFGtri (BA45, $x, y, z = \pm 48, 27, 12$) and MTG ($x, y, z = \pm 48, -60, 18$) for the ventral pathway. We placed VOIs in the medial temporal lobe ($x, y, z = \pm 40, -20, -7$) and temporo-occipital junction ($x, y, z = \pm 35, -68, -6$) for ILF according to a previous study (Jou et al., 2011). We performed PDT within each hemisphere: IFGop-caudate nucleus, IFGop-STG/SMG (dorsal pathway), IFGtri-MTG (ventral pathway), and ILF (eight tracts in total). The voxel values indexing probabilistic strength of connectivity across the two regions were averaged within each PDT pathway and then subjected to logarithmic transformation, producing a parameter for statistical analyses (hereafter called the connectivity parameter). Statistical tests on the connectivity parameter computed from PDT were performed using SPSS 17.0 (IBM).

In the cohort study, we performed a 2-by-2 mixed repeated-measures ANOVA with time as a within-subject variable and the group as a between-subject variable. In addition, a correlation analysis was performed to test if the training-related changes in TOEIC score were correlated with those of the connectivity parameter from the eight tracts: bilateral IFGop-caudate pathway, dorsal pathway, ventral pathway, and ILF. Moreover, to test the possibility that connectivity of the specific tracts before intervention could predict the degree of improvement in L2 ability, a correlation analysis was performed between the connectivity parameter of the PDT in the Pre condition and the improvement of L2 proficiency.

Analysis of follow-up behavioral and imaging data. Only in the training group, for which follow-up data were available, TOEIC score, GM volume, FA value, and the strength of connectivity of the eight specific tracts were compared across the Pre, Post-1, and Post-2 conditions. To retrieve the GM volume and FA values of interest, we applied spherical VOIs with a 5 mm radius to the right IFGop and WM beneath the right IFG based on the results from the cross-sectional study. The data were fed into a one-way repeated-measures ANOVA, followed by *post hoc* comparisons with Tukey's HSD test. Finally, we ran a correlation analysis between the GM changes from the Pre to Post-1 or Post-1 to Post-2 and FA changes from the Pre to Post-1 or Post-1 to Post-2. This analysis suggested the existence of different subgroups following the completion of training. Indeed, three participants continued L2 learning of their own motivation after our intervention, while the rest of the participants ($n = 21$) did not. Hence, a subgroup analysis was conducted to compare the changes of TOEIC, GM, or FA from Post-1 to Post-2 between the two groups. Because one of the groups consisted of a small number of participants, we performed this test with a nonparametric, Mann–Whitney *U* test.

Results
Cross-sectional study
Mean EVT and NART scores were 20.6 and 41.8, respectively (Table 1). The large variance of these values supported that L2 proficiency of the participants was variable.

The VBM analysis showed that individuals with more extensive L2 vocabulary (EVT) had significantly larger GM volume in the IFGop corresponding to Brodmann area (BA) 44, caudate nuclei, STG/SMG, and anterior cingulate cortex, all bilaterally (Fig. 1, Table 2). The correlation of the EVT score and GM volume was found more strongly in the right hemisphere for the IFGop (LI = −0.40) and within the range of symmetry for the caudate nucleus (LI = −0.15) and STG (LI = −0.17). The TBSS analysis showed that individuals with richer L2 vocabulary showed higher FA values in the subcortical WM beneath the IFGop (sub-IFGop), ILF, and AF only in the right hemisphere. The PDT identified bilateral IFGop-caudate nucleus and IFGop-STG/SMG (dorsal pathway) in all individuals. A statistical analysis of the connectivity parameters retrieved from the four tracts revealed correlation of the EV score with connectivity of the IFGop-caudate nucleus ($p < 0.001$) and the IFGop-STG/SMG (dorsal pathway) ($p < 0.001$) in the right hemisphere, but not in

Table 2. Significant correlation between EVT score and GM volume ($p < 0.05$ corrected) in the cross-sectional study

Anatomical location	Coordinates			Z-value	P corrected
	x	y	z		
Right IGFtri (BA44)	36	11	28	7.84	0.00
Right caudate nucleus	12	8	19	5.92	0.00
Right STG/SMG	56	−31	21	5.42	0.00
Right MTG	49	−20	0	4.08	0.00
Left IFG (BA44, 45)	−38	16	30	6.69	0.00
Left superior frontal gyrus	−22	35	45	5.45	0.00
Left caudate nucleus	−4	16	−2	5.26	0.00
Left STG/SMG	−48	−50	26	4.92	0.00

The coordinates (x, y, z) indicate local maxima in each brain region according to the MNI template.

the left hemisphere (IFGop-caudate, $p = 0.09$; dorsal pathway, $p = 0.11$; Fig. 1).

NART score did not correlate with GM volume (VBM analysis) or FA value (TBSS analysis) at the predetermined statistical threshold. When a lenient threshold (uncorrected $p < 0.001$) was applied to the VBM analysis, we detected a trend toward correlation between NART score and GM volume in the middle temporal and postcentral gyri in the left hemisphere.

Cohort study: behavioral data
Baseline profiles did not differ between the TG and CG participants in terms of age, sex, IQs, personality traits, or L2 ability (Table 1).

The total learning time, recorded on log files, was 45.5 h (SD = 3.6, range 25.1 − 64.6) across the whole training period. Just after L2 training (Post-1), the TG showed a 29 ± 21.3% (mean ± SD) improvement in total TOEIC score, whereas the CG group showed no change (0 ± 8.1%; Fig. 2a), revealing the significant effect of L2 training ($F_{(3,82)} = 27.15$, $p < 0.001$ by mixed repeated-measures ANOVA; Table 3). Specifically, improvement was found in the listening ($F_{(3,82)} = 17.64$, $p < 0.001$) and reading ($F_{(3,82)} = 10.15$, $p = 0.001$) sections of TOEIC, but not in the grammar section ($F_{(3,82)} = 0.50$, $p = 0.42$; mixed repeated-measures ANOVA). This finding is reasonable because the training program included items for vocabulary and listening (i.e., pronunciation), but not for grammar. The changes in TOEIC score paralleled those of the EVT score ($r = 0.41$, $p = 0.03$), whereas no correlation was found with those of the NART score ($r = 0.15$, $p = 0.53$).

Follow-up L2 proficiency tests were obtained in the TG participants a year after the training program. One-way repeated-measures ANOVA showed significant differences in TOEIC score across the three time points ($F_{(1.8, 38.6)} = 24.96$, $p < 0.001$ by repeated-measures ANOVA with sphericity correction). Post hoc comparisons (Tukey's HSD test) demonstrated significant increases in score from the Pre to Post-1 ($p < 0.001$) conditions and significant decreases from Post-1 to Post-2 ($p = 0.02$), resulting in no difference in score between Pre and Post-2.
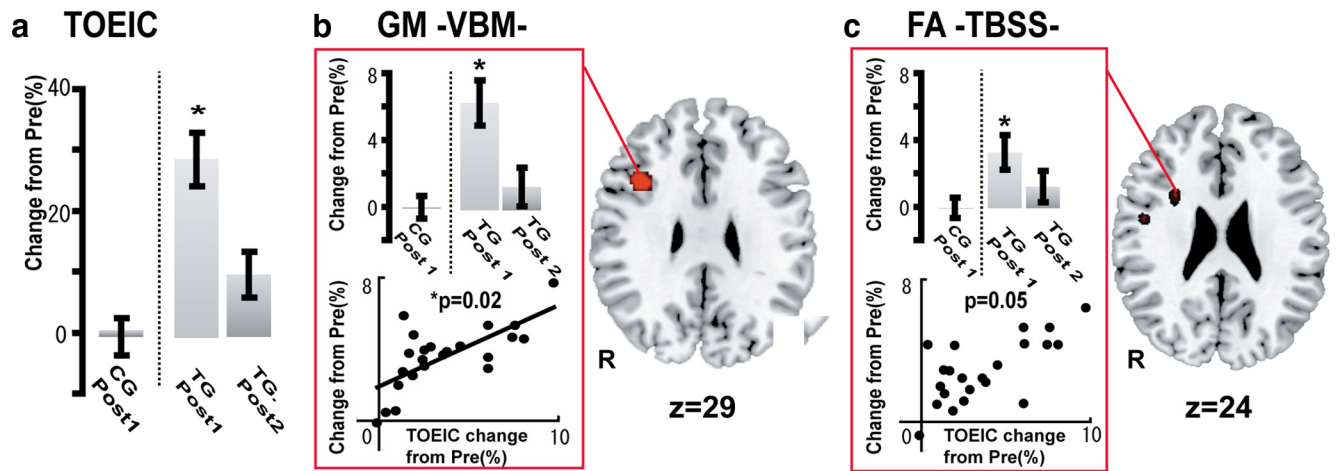
Cohort study: imaging data
Before the training (Pre), no significant differences in brain structure were detected in the GM-VBM, WM-TBSS, and connectivity of the eight specific tracts (bilateral IFGop-caudate, dorsal pathway, ventral pathway, and ILF) between the TG and the CG, supporting the homogeneity of the two groups. Further, we examined whether brain architecture at the Pre stage could predict L2 ability improvement gains after the training program in the TG. However, we failed to find significant correlations between the training-induced improvement of TOEIC scores

and Pre-GM volume (VBM), Pre-FA values (TBSS), or Pre-connectivity of the eight tracts (PDT).

The analyses of the MRI data (Post-1 vs Pre) identified training-induced increases in both GM and WM of the right hemisphere for the TG compared with the CG (significant time-by-group interaction). In the TG compared with the CG, VBM analysis identified significant increases in GM volume only in the right IFGop (5.1 ± 2.3%), corresponding to the neural outcome of plastic changes induced by the L2 vocabulary learning (Fig. 2b). The training-induced increases of GM volume in the right IFGop were correlated with those of TOEIC score ($r = 0.52$, $p = 0.02$).

In the TG, FA values increased from Pre to Post-1 in WM in the right sub-IFGop region (3.6 ± 1.6%; Fig. 2c). This increase in FA showed a strong tendency to correlate with the increase in TOEIC score ($r = 0.33$, $p = 0.05$). Using the PDT analyses, we further discovered learning-related increases in the connectivity parameter of the IFGop-caudate head pathway (4.5% ± 6.4, $F_{(3,82)} = 7.83$, $p = 0.01$) and that of the dorsal pathway (3.7 ± 6.4%, $F_{(3,82)} = 6.72$, $p = 0.01$) in the right hemisphere (2-by-2 repeated-measures ANOVA). The homologous tracts in the left hemisphere did not show the corresponding changes (2.5 ± 4.4%, $F_{(3,82)} = 0.16$, $p = 0.69$ for the left IFGop-caudate connectivity; 2.7 ± 5.2%, $F_{(3,82)} = 0.16$, $p = 0.65$ for the left dorsal pathway connectivity; Fig. 3). There were no significant L2 training-induced connectivity changes in the ventral pathway (2.1 ± 4.0%, $F_{(3,82)} = 0.16$, $p = 0.69$ for the right; 1.9 ± 4.5%, $F_{(3,82)} = 0.26$, $p = 0.70$ for the left) or in the ILF (1.8 ± 5.0%, $F_{(3,82)} = 0.22$, $p = 0.41$ for the right; 2.0 ± 5.1%, $F_{(3,82)} = 0.31$, $p = 0.49$ for the left). The training-induced increases in TOEIC score were correlated with those of the right IFGop-caudate connectivity ($p = 0.02$, $r = 0.59$), but not of the left IFGop-caudate ($p = 0.13$, $r = 0.25$), right dorsal pathway ($p = 0.07$, $r = 0.29$), left dorsal pathway ($p = 0.30$, $r = 0.14$), right ILF ($p = 0.14$, $r = 0.19$), or left ILF connectivity ($p = 0.25$, $r = 0.20$; Fig. 3a,e,f,j,k,o).

We re-examined L2 proficiency and brain structure of the participants in the TG a year after the completion of the training program. One-way repeated-measures ANOVA showed significant differences in GM ($F_{(1.8, 38.6)} = 477.0$, $p < 0.001$), FA ($F_{(1.8, 38.6)} = 233.8$, $p < 0.001$), and right IFGop-caudate connectivity ($F_{(1.8, 38.6)} = 331.0$, $p < 0.001$) across the three time points. Consistent with the aforementioned group comparison, post hoc comparisons (Tukey's HSD test) between Pre and Post-1 demonstrated significant increases in GM volume in the right IFG ($p < 0.001$), FA beneath the right IFGop ($p = 0.03$), and connectivity between the right IFGop and caudate head ($p = 0.02$). Furthermore, the increment of GM volume and FA values from Pre to Post-1 were significantly correlated (Fig. 4a), suggesting parallel plastic changes in GM and WM. Nevertheless, post hoc comparisons (Tukey's HSD test) between Post-1 and Post-2 showed significant decreases in GM volume in the right IFGop (−4.3 ± 2.5%, $p = 0.04$) and FA beneath the right IFG (−3.7 ± 4.2%, $p = 0.04$). In addition, the decreases of GM volume from Post-1 to Post-2 were significantly correlated with those of FA values ($p < 0.001$, $r = 0.65$; Fig. 4b). That means the parallelism of the "negative" changes in GM and WM. We also noticed a non-negligible trend toward decreases in right IFGop- caudate head connectivity (−2.3 ± 8.1%, $p = 0.05$) in the connectivity of the dorsal pathway (−2.0 ± 7.2%, $p = 0.06$). Resultantly, no difference was found between the Pre and Post-2 values for any of the imaging parameters. These findings indicate that in most of the participants, training-induced behavioral values and macroscopic neural structure returned to the pretraining state, suggesting a phenomenon of neural "elasticity." Through the assessment of the

Figure 2. Results from the cohort study. Shown are changes in L2 competence (TOEIC), GM, and FA in the cohort study; *$p < 0.05$. There were no significant differences in behavioral or imaging parameters between the TG and CG at the Pre stage. *a*, Mean total TOEIC score changes from Pre to Post-1 in CG (left of the dotted line) and from Pre to Post-1 and Post-2 in TG (right of the dotted line). Error bars indicate standard error of the mean. Asterisks show significant differences with $p < 0.05$. *b*, Right, Colored voxels represent a cluster in the right IFGop ($x, y, z = 36, 11, 28$; z value = 7.84) showing significant GM increases from Pre to Post-1 in TG compared with those in CG (group-by-time interaction, $p < 0.05$ FWE-corrected). Mean GM volume changes are shown (upper left) for the right IFGop from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. Changes in GM volume for the right IFGop significantly correlated with improvement in total TOEIC score across individuals (lower left). *c*, Right, Colored voxels represent WM clusters showing significant FA increases from Pre to Post-1 in TG than those in CG (group-by-time interaction, $p < 0.05$ FWE-corrected). Upper left, Shows mean FA changes from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. The change in FA value beneath the right IFGop was significantly correlated with change in total TOEIC score ($p < 0.05$, lower left). R, right.

Table 3. Means and SDs of subject L2 proficiency tests

	TG (n = 24)			CG (n = 20)		
	Pre	Post-1	Post-2	Pre	Post-1	Post-2
TOEIC (200)	102.5 ± 4.0	129.6 ± 5.1	114.1 ± 4.5	100.8 ± 4.3	103.7 ± 4.3	—
EVT (100)	11.3 ± 10.9	16.9 ± 8.1	12.4 ± 11.0	12.9 ± 12.1	12.0 ± 10.1	—
NART (100)	35.6 ± 4.5	37.8 ± 4.6	37.0 ± 4.0	39.8 ± 4.0	38.0 ± 4.4	—

Post-2 data, however, we noticed that three participants showed further improvement in L2 ability from Post-1 to Post-2 (8.3, 7.5, and 7.8% increases in TOEIC score), despite the overall decreases of the TOEIC score as a group. Two of those exceptional participants continued to study L2 to pass a certification examination, and one participated in a short-term study abroad program after finishing our training program. Even though these were anecdotal findings in a small number of participants, we deemed them important and performed a subgroup analysis: three participants with increased TOIEC score from Post-1 to Post-2 (continued subgroup) and the rest ($n = 21$; discontinued subgroup). There were significant differences in changes of TOEIC score from Post-1 to Post-2 between the continued and discontinued (mean $-21\% \pm 8.5$) subgroups (Mann–Whitney U test, $p < 0.001$). Intriguingly, the continued group showed increases in GM volume in the right IFGop (3.1, 2.9, and 2.5%, respectively) from Post-1 to Post-2 as opposed to the discontinued group (mean $-4.5\% \pm$ SD 1.2). The same held true for the FA value in the right sub-IFGop (2.8, 3.2, and 2.1%, respectively, for the continued group) as opposed to the discontinued group (mean $-3.6\% \pm 1.5$; Fig. 4*b*). These differences reached significance for the GM change ($p = 0.001$ by Mann–Whitney U test) and for the FA change ($p = 0.001$ by Mann–Whitney U test).

Discussion

Currently, international communication relies heavily on English. This fact imposes educational challenges to non-English native cultures in the rapidly globalizing world, prompting us to advance our understanding of the mechanisms of L2 acquisition. The human capacity for language relies on neural networks that orchestrate lexicosemantic, phonological, and syntactic subsystems, which should undergo adequate tuning for L2 learning. However, nonsyntactic aspects of L2 learning are scarcely understood hitherto. Here we focused on L2 vocabulary learning, which should involve the lexicosemantic and phonological subsystems. The present cross-sectional study showed that bilateral front-subcortical-parietotemporal areas, predominantly in the right hemisphere, might underlie superior L2 vocabulary ability in late L2 learners. The cohort experiment showed that the brain architecture before learning could not predict the future gain in L2 ability after learning. This finding argues against the idea that privileged persons with developed language network have advantages in L2 learning, although it holds true for discriminating speech sounds of unfamiliar L2 (Golestani and Pallier, 2007; Golestani et al., 2007). This discrepancy suggests that relative importance of predispositions and environmental effects may differ, depending on which aspects of L2 should be learned. The cohort experiment clearly disclosed that the gains and losses of L2 ability correlated with increases and decreases, respectively, of GM structure and WM connectivity. This longitudinal consistency of the behavior–structure relationship within individuals highlights the notion that use-dependent plastic changes of neural networks underlie L2 learning. No overall gains of TOEIC score in the Post-2 suggested that the present e-training protocol was not sufficiently effective in leaving a long-term signature of L2 vocabulary learning. Note, however, that the loss of L2 ability measured by the behavioral test does not necessarily mean that the once-remembered items had completely disappeared from long-term memory. It is also possible that the items could not be retrieved in a contextually timely manner. Such presumable mechanisms agree with reduction of connectivity (i.e., disconnection) between the frontal executive areas and temporoparietal cortices.

Recent imaging studies have begun to show GM/WM changes related to learning. Biological interpretations of those imaging measures are still challenging. Many possible mechanisms could be involved: neurogenesis, gliogenesis, synaptogenesis, and vas-
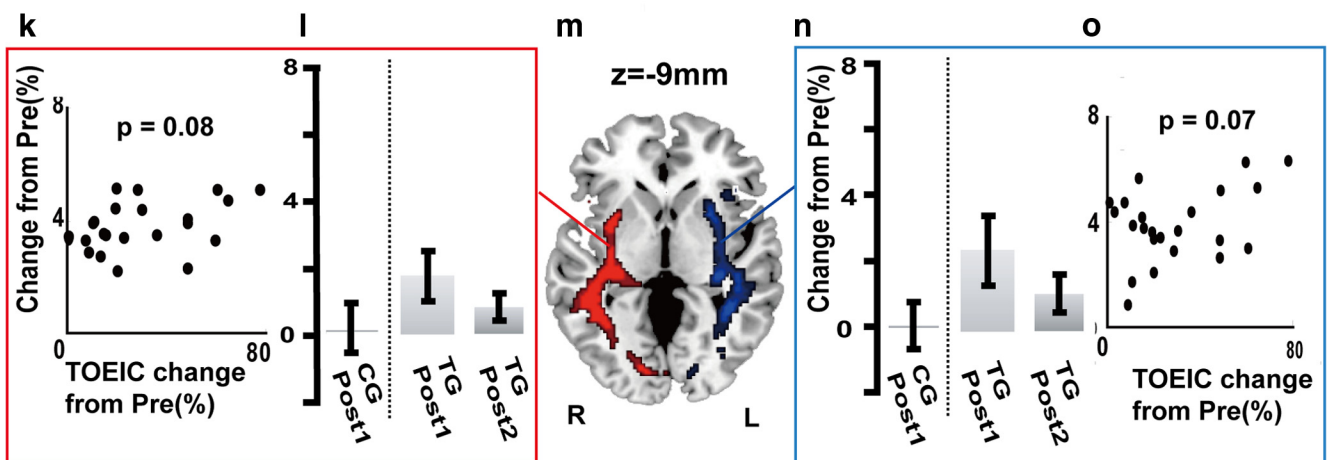
IFG-CN connectivity
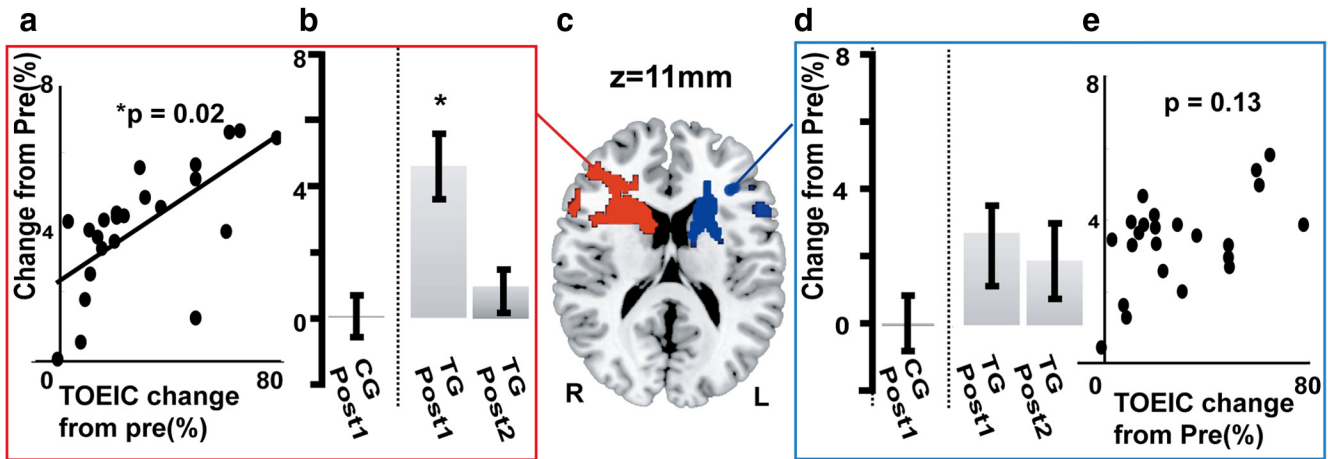
IFGop-STG/SMG connectivity (dorsal pathway)
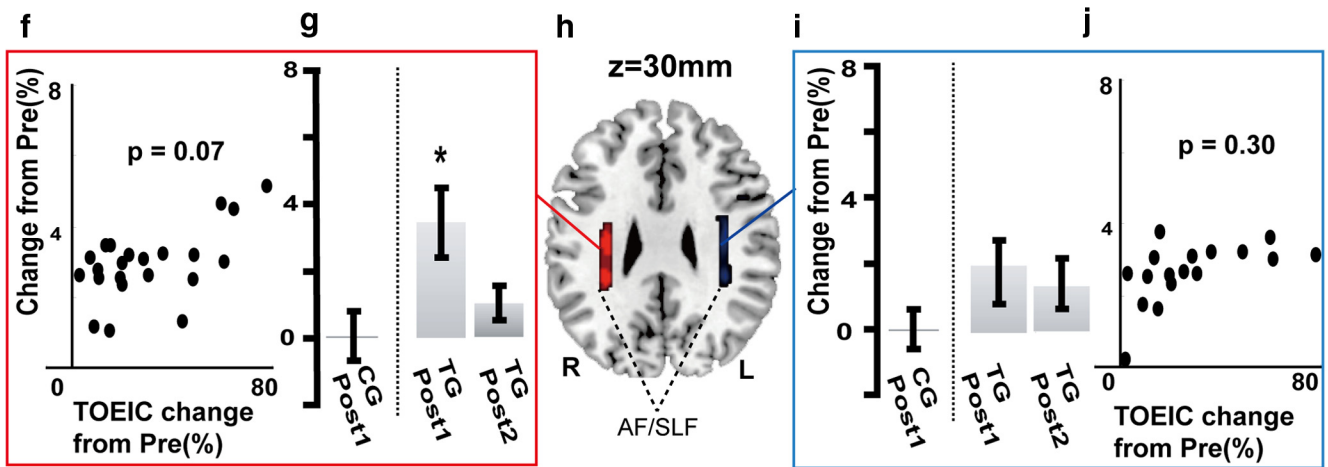
# IFGtri-MTG connectivity (ventral pathway)

Figure 3. PDT results from cohort study. The strength of connectivity parameters for the IFGop-caudate pathway and IFGop-STG/SMG tract (dorsal pathway) were increased after the training period in the right hemisphere, but not in the left. The improvement in L2 competence (TOEIC) was correlated with the increase in the right IFGop-caudate connectivity parameter only. Error bars indicate standard error of the mean. Asterisk shows significant differences with $p < 0.05$. *a*, The change in the right IFGop-caudate connectivity parameter positively correlated with improvement in the total TOEIC score $p < 0.05$. *b*, Mean changes in right IFGop-caudate connectivity parameter from Pre to Post-1 in CG (left of the dotted line) and from Pre to Post-1 and Post-2 in TG (right of the dotted line). *c*, The red and blue areas show group-averaged IFGop-caudate tracts in the right and left hemispheres, respectively (data from Post-1 in TG). The seed and target regions were set in the right IFGop and the whole caudate nucleus (CN), respectively. *d*, Shown are left IFGop-caudate connectivity parameter changes from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *e*, The changes in the left IFGop-caudate connectivity parameter did not correlate with improvement in the total TOEIC score. *f*, Changes of connectivity parameter in the right dorsal pathway (IFGop-STG/SMG) were not significantly correlated with improvement in the total TOEIC score. *g*, Shown are changes in the right dorsal pathway connectivity parameter (*Figure legend continues.*)

cular changes for GM, and remodeling of the myelin sheath and activity-dependent axonal changes for WM reorganization indexed by FA (Scholz et al., 2009; Zatorre et al., 2012). Recently, an animal experiment associated increases in FA with increased expression of a marker of myelination (Blumenfeld-Katzir et al., 2011). An important finding here was that the bidirectional changes of the behavioral, GM, and WM parameters paralleled each other, in contrast to a previous study showing discrepancy between training-induced changes in GM and WM (Taubert et al., 2010). The coupling of behavioral, GM, and WM changes supports that learning depends upon experience-dependent activities correlated across connected regions (Fields, 2005; Zatorre et al., 2012).

The present study has provided the first compelling evidence for increased connectivity of specific tracts underlying L2 ability and acquisition. We found that the IFGop constituted networks with important language-related nodes: the STG/SMG (Price et al., 1999; Price and Crinion, 2005; Saur et al., 2008; Carreiras et al., 2009) and the caudate nucleus (Crinion et al., 2006; Friederici, 2006). The IFGop and caudate nucleus probably constitute the corticobasal ganglia circuits relevant to language processing and learning. Impairments of the caudate nucleus are identified in people with mutation of FOXP2, which is an important genetic predisposition for language acquisition (Liégeois et al., 2003; Enard et al., 2009). Additionally, caudate activity underlies lexicosemantic control in bilinguals (Crinion et al., 2006), and reduced striatal dopamine releases impair language processing (Tettamanti et al., 2005). The corticobasal ganglia circuits are suggested for reward-based reinforcement learning since they receive both contextual information from the cortex and reward signals from dopaminergic neurons (Doya, 2008). Reinforcement learning might play a part in enhancing executive control of the IFG over the mechanisms for acquiring L2 vocabulary.

We found increased connectivity of the "dorsal pathway," corresponding mainly to the temporal part of AF involved in phonological processing (Rilling et al., 2008). AF underlies successful associative learning between sounds and words (Wong et al., 2011; Yeatman et al., 2011). AF may thus integrate phonological and se-
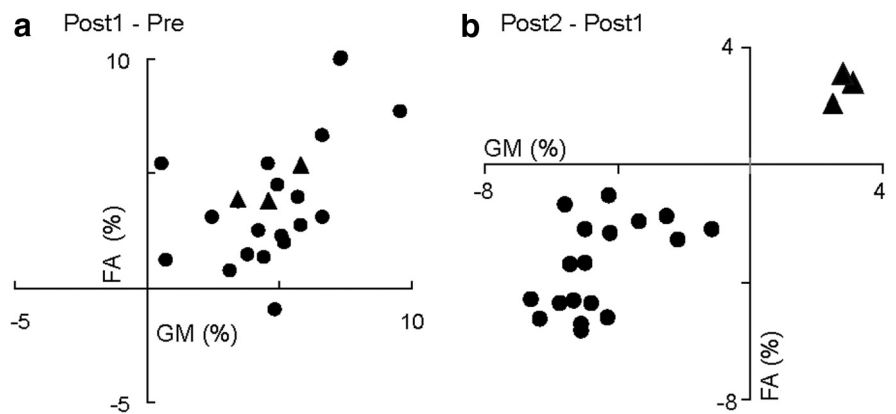
Figure 4. Correlation between GM and WM changes induced by L2 training. *a*, Changes in GM volume (abscissa) and FA (ordinate) from Pre to Post-1 in TG. The plots show significant correlation between GM volume and FA changes ($p = 0.03$). *b*, Changes in GM volume and FA from Post-1 to Post-2. The plots show significant correlation between GM volume change and FA change ($p < 0.001$, $r = 0.65$). The circles indicate participants with decreased TOEIC scores and the triangles indicate subjects with increased TOEIC scores from Post-1 to Post-2. The filled circles indicate participants who showed decrement of the total TOEIC scores from Post-1 to Post-2 while the filled triangles indicate participants who showed increment of the total TOEIC scores for the same period (see main text for the subgroup analysis).

(*Figure legend continued.*) from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. Error bars indicate SEM. Asterisks show significance differences with $p < 0.05$. *h*, The red and blue areas represent the group-averaged dorsal pathway tractography including AF and SLF in the right and left hemispheres, respectively (data from Post-1 in TG). *i*, Mean changes in the left dorsal pathway connectivity parameter from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *j*, Changes in the left dorsal pathway connectivity parameter did not correlate with the change in the total TOEIC score. *k*, Changes in the right ventral pathway (IFGtri-MTG) parameter were not significantly correlated with improvement in the total TOEIC score. *l*, Changes in the right ventral pathway connectivity parameter from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *m*, The red and blue areas represent the group-averaged ventral pathway tractography in the right and left hemispheres, respectively (data from Post-1 in TG). *n*, Mean changes in the left ventral pathway connectivity parameter from Pre to Post-1 in CG and from Pre to Post-1 and Post-2 in TG. *o*, Changes in the left ventral pathway connectivity parameter did not correlate with improvement in the total TOEIC score. The connectivity parameters for the inferior longitudinal fascicles did not change in either hemisphere (data not shown). R, right; L, left.

mantic aspects of L2 vocabulary. The predominant involvement of the phonological subsystem in L2 vocabulary learning agrees with the present finding showing little connectivity changes in the ventral pathway and ILF, which are involved mainly in semantic processing (Saur et al., 2008). The predominant involvement of the dorsal pathway suggests that phonology learning may have played a pivotal role in the present L2 training paradigm. A future study, however, will be needed to dissect the substrates for the lexicosemantic and phonological aspects of L2 vocabulary learning.

A rather surprising outcome was that the IFGop-caudate-STG/SMG network correlated with L2 vocabulary competence and its learning-induced plastic changes were lateralized to the right hemisphere. These observations appear to contradict with mounting evidence indicating significance of the left hemispheric language areas for L2 (Tettamanti et al., 2002; Musso et al., 2003; Abutalebi and Green, 2007; Abutalebi et al., 2008; Sakai et al., 2009; Mårtensson et al., 2012; Schlegel et al., 2012; Ghazi Saidi et al., 2013). Moreover, learning of new languages induces plastic changes in GM and WM in the left hemisphere (Schlegel et al., 2012). In early learning stages, however, the right frontal cortex is involved in acquisition of artificial grammar and natural language (Fletcher et al., 1999, 2005; Tettamanti et al., 2002; Musso et al., 2003). Recent evidence has begun to highlight the importance of the right hemisphere for vocabulary and phonological aspects of L2 processing. The right STG/SMG is involved in non-L1 vocabulary learning (Smith et al., 2006; Jeong et al., 2010; Raboyeau et al., 2010; Veroude et al., 2010). We previously showed that right IFGop activity for switching phonology from L1 to L2 was correlated with L2 vocabulary levels (Hosoda et al., 2012). Others showed enhanced activity in the right prefrontal areas during picture naming in L2 (Videsott et al., 2010), coupling of right IFG activity with proficiency of word production in L2 (Calabrese et al., 2001; Vingerhoets et al., 2003; van Ettinger-Veenstra et al., 2010), and involvement of the right hemisphere in the control of verbal interference in bilinguals (Filippi et al., 2011). Moreover, a case study suggested an essential role of the right IFG for L2 (April and Tse, 1977); a dextral late bilingual patient suffering from a right IFG lesion had severe difficulty in finding words and reading more in L2 than in L1. Hence, evidence certainly supports the roles of the right IFG and STG in nongrammatical aspects of L2 usage, although they may not be as well recognized as the correlates of L2 compared with the left counterparts.

A limitation of the cohort experiment was no behavioral control over CG during the training period. Although we cannot completely exclude the effects of nonlanguage e-learning factors on the training-induced changes, we considered them unlikely as the sole explanation of the learning-induced plastic changes because of the following reasons. First, there was no correlation of the total e-training time with the GM/WM changes. Second, previous long-term computer-learning studies did not report changes of right IFG, caudate nucleus, or STG (Takeuchi et al., 2010; Wan et al., 2012). Consistently, preliminary results from our computer-based sequence learning experiment for 10 weeks failed to find changes in the right IFGop, caudate nucleus, or STG (C. Hosoda, M. Honda, T. Hanakawa, unpublished observation).

Paucity of evidence allows us only to speculate specific functions of the right IFGop in L2 vocabulary learning. We theorize that the right IFGop might link lexicosemantic-phonological knowledge between L1 and L2. For late L2 vocabulary learning, one should associate new L2 vocabulary temporarily represented in short-term memory with existing L1 vocabulary represented in the left hemisphere. This agrees with a finding that acquisition of new L2 phonology recruits the left-lateralized language network (Paulesu et al., 2009). However, more long-term encoding-retrieval processes may modify bilateral networks since the right prefrontal areas are important for accessing to long-term memory (Ranganath et al., 2007). Another simple, yet plausible, explanation is a spillover of language representations/control functions from the left to the right hemisphere. This could be particularly important here since Japanese is one of the most linguistically distant languages from English (Chiswick and Miller, 2005). The spillover concept is consistent with findings that mathematics/arithmetic experts show greater right hemispheric activity than control subjects for whom left hemispheric activity is more relevant (Hanakawa et al., 2003; Aydin et al., 2007). These findings may characterize a repertoire of use-dependent dynamic reorganization of the brain. A dramatic example is a child who shifted the originally left hemispheric language centers to the mirror-reversed, right hemispheric sites after surgical removal of the left hemisphere (Hertz-Pannier et al., 2002). Language experience-dependent changes may likely occur in the right IFG because of weaker genetic influences on the right IFG compared with the left (Thompson et al., 2001).

In conclusion, the present study indicates that the macroscopic reorganization of the IFGop-caudate-STG network, especially on the right, can underlie L2 vocabulary learning in adults (1500/1500 words).

References

Abutalebi J, Green DW (2007) Bilingual language production: the neurocognition of language representation and control. J Neurolinguistics 20:242–275. CrossRef

Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, Lazeyras F, Cappa SF, Khateb A (2008) Language control and lexical competition in bilinguals: an event-related FMRI study. Cereb Cortex 18:1496–1505. Medline

Alexander AL, Lee JE, Lazar M, Field AS (2007) Diffusion tensor imaging of the brain. Neurotherapeutics 4:316–329. CrossRef Medline

April RS, Tse PC (1977) Cross aphasia in a Chinese bilingual dextral. Arch Neurol 34:766–770. CrossRef Medline

Ashburner J (2007) A fast diffeomorphic image registration algorithm. Neuroimage 38:95–113. CrossRef Medline

Aydin K, Ucar A, Oguz KK, Okur OO, Agayev A, Unal Z, Yilmaz S, Ozturk C (2007) Increased gray matter density in the parietal cortex of mathematicians: a voxel-based morphometry study. AJNR Am J Neuroradiol 28:1859–1864. CrossRef Medline

Beaulieu C (2002) The basis of anisotropic water diffusion in the nervous system–a technical review. NMR Biomed 15:435–455. CrossRef Medline

Blumenfeld-Katzir T, Pasternak O, Dagan M, Assaf Y (2011) Diffusion MRI of structural brain plasticity induced by a learning and memory task. PLoS One 6:e20678. CrossRef Medline

Briellmann RS, Mitchell LA, Waites AB, Abbott DF, Pell GS, Saling MM, Jackson GD (2003) Correlation between language organization and diffusion tensor abnormalities in refractory partial epilepsy. Epilepsia 44:1541–1545. CrossRef Medline

Calabrese P, Neufeld H, Falk A, Markowitsch HJ, Müller C, Heuser L, Gehlen W, Durwen HF (2001) Wortgenerierung bei Bilingualen-eine fMRT-Studie mit Implikationen für Sprach-und Gedächtnisprozesse. Fortschr Neurol Psychiatr 69:42–49. CrossRef Medline

Carreiras M, Seghier ML, Baquero S, Estévez A, Lozano A, Devlin JT, Price CJ (2009) An anatomical signature for literacy. Nature 461:983–986. CrossRef Medline

Catani M, Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. Cortex 44:953–961. CrossRef Medline

Chiswick BR, Miller PW (2005) Linguistic distance: a quantitative measure of the distance between English and other languages. J Multiling Multicult Dev 26:1–11. CrossRef

Coady J, Huckin T (1997) Second Language Vocabulary Acquisition: Cambridge UP.

Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, Fukuyama H, Stockton K, Usui K, Green DW, Price CJ (2006) Language control in the bilingual brain. Science 312:1537–1540. CrossRef Medline

Dowens MG, Vergara M, Barber HA, Carreiras M (2010) Morphosyntactic processing in late second-language learners. J Cogn Neurosci 22:1870–1887. CrossRef Medline

Doya K (2008) Modulators of decision making. Nat Neurosci 11:410–416. CrossRef Medline

Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A (2004) Neuroplasticity: changes in grey matter induced by training. Nature 427:311–312. CrossRef Medline

Ellis R, Heimbach R, Tanaka Y, Yamazaki A (1999) Learning a second language through interaction. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, et al. (2009) A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. Cell 137:961–971. CrossRef Medline

Fields RD (2005) Myelination: an overlooked mechanism of synaptic plasticity? Neuroscientist 11:528–531. CrossRef Medline

Filippi R, Richardson FM, Dick F, Leech R, Green DW, Thomas MS, Price CJ (2011) The right posterior paravermis and the control of language interference. J Neurosci 31:10732–10740. CrossRef Medline

Fletcher P, Büchel C, Josephs O, Friston K, Dolan R (1999) Learning-related neuronal responses in prefrontal cortex studied with functional neuroimaging. Cereb Cortex 9:168–178. CrossRef Medline

Fletcher PC, Zafiris O, Frith CD, Honey RA, Corlett PR, Zilles K, Fink GR (2005) On the benefits of not trying: brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. Cereb Cortex 15:1002–1015. Medline

Flöel A, de Vries MH, Scholz J, Breitenstein C, Johansen-Berg H (2009) White matter integrity in the vicinity of Broca's area predicts grammar learning success. Neuroimage 47:1974–1981. CrossRef Medline

Friederici AD (2006) What's in control of language? Nat Neurosci 9:991–992. CrossRef Medline

Ghazi Saidi L, Perlbarg V, Marrelec G, Pélégrini-Issac M, Benali H, Ansaldo AI (2013) Functional connectivity changes in second language vocabulary learning. Brain Lang 124:56–65. CrossRef Medline

Golestani N, Pallier C (2007) Anatomical correlates of foreign speech sound production. Cereb Cortex 17:929–934. Medline

Golestani N, Molko N, Dehaene S, LeBihan D, Pallier C (2007) Brain structure predicts the learning of foreign speech sounds. Cereb Cortex 17:575–582. Medline

Grogan A, Green DW, Ali N, Crinion JT, Price CJ (2009) Structural correlates of semantic and phonemic fluency ability in first and second languages. Cereb Cortex 19:2690–2698. CrossRef Medline

Grogan A, Parker Jones Ö, Ali N, Crinion J, Orabona S, Mechias ML, Ramsden S, Green DW, Price CJ (2012) Structural correlates for lexical effi-

ciency and number of languages in non-native speakers of English. Neuropsychologia 50:1347–1352. CrossRef Medline

Hanakawa T, Honda M, Okada T, Fukuyama H, Shibasaki H (2003) Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. Neuroimage 19:296–307. CrossRef Medline

Hernandez AE, Dapretto M, Mazziotta J, Bookheimer S (2001) Language switching and language representation in Spanish-English bilinguals: an fMRI study. Neuroimage 14:510–520. CrossRef Medline

Hertz-Pannier L, Chiron C, Jambaqué I, Renaux-Kieffer V, Van de Moortele PF, Delalande O, Fohlen M, Brunelle F, Le Bihan D (2002) Late plasticity for language in a child's non-dominant hemisphere: a pre- and post-surgery fMRI study. Brain 125:361–372. CrossRef Medline

Hosoda C, Hanakawa T, Nariai T, Ohno K, Honda M (2012) Neural mechanisms of language switch. J Neurolinguistics 25:44–61. CrossRef

Hsu AS, Chater N, Vitányi PM (2011) The probabilistic analysis of language acquisition: theoretical, computational, and experimental analysis. Cognition 120:380–390. Medline

Jeong H, Sugiura M, Sassa Y, Wakusawa K, Horie K, Sato S, Kawashima R (2010) Learning second language vocabulary: neural dissociation of situation-based learning and text-based learning. Neuroimage 50:802–809. CrossRef Medline

Johansen-Berg H (2007) Structural plasticity: rewiring the brain. Curr Biol 17:R141–R144. CrossRef Medline

Johansen-Berg H (2012) The future of functionally-related structural change assessment. Neuroimage 62:1293–1298. CrossRef Medline

Johansen-Berg H, Baptista CS, Thomas AG (2012) Human structural plasticity at record speed. Neuron 73:1058–1060. CrossRef Medline

Jou RJ, Jackowski AP, Papademetris X, Rajeevan N, Staib LH, Volkmar FR (2011) Diffusion tensor imaging in autism spectrum disorders: preliminary evidence of abnormal neural connectivity. Aust N Z J Psychiatry 45:153–162. CrossRef Medline

Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. Science 303:1023–1026. CrossRef Medline

Kuhl PK (2010) Brain mechanisms in early language acquisition. Neuron 67:713–727. CrossRef Medline

Laufer B (1989) A factor of difficulty in vocabulary learning: deceptive transparency. AILA Rev 6:10–20.

Le Bihan D, Johansen-Berg H (2012) Diffusion MRI at 25: exploring brain tissue structure and function. Neuroimage 61:324–341. Medline

Liégeois F, Baldeweg T, Connelly A, Gadian DG, Mishkin M, Vargha-Khadem F (2003) Language fMRI abnormalities associated with FOXP2 gene mutation. Nat Neurosci 6:1230–1237. CrossRef Medline

Mårtensson J, Eriksson J, Bodammer NC, Lindgren M, Johansson M, Nyberg L, Lövdén M (2012) Growth of language-related brain areas after foreign language learning. Neuroimage 63:240–244. CrossRef Medline

Mechelli A, Crinion JT, Noppeney U, O'Doherty J, Ashburner J, Frackowiak RS, Price CJ (2004) Neurolinguistics: structural plasticity in the bilingual brain. Nature 431:757. CrossRef Medline

Meschyan GaH, A (2002) Is native-language decoding skill related to second-language learning? J Educ Psychol 94:14–22. CrossRef

Musso M, Moro A, Glauche V, Rijntjes M, Reichenbach J, Büchel C, Weiller C (2003) Broca's area and the language instinct. Nat Neurosci 6:774–781. CrossRef Medline

Paulesu E, Vallar G, Berlingeri M, Signorini M, Vitali P, Burani C, Perani D, Fazio F (2009) Supercalifragilisticexpialidocious: how the brain learns words never heard before. Neuroimage 45:1368–1377. CrossRef Medline

Peltier J, Nicot B, Baroncini M, Zunon-Kipré Y, Haidara A, Havet E, Foulon P, Page C, Lejeune JP, Le Gars D (2011) [Anatomy of the periventricular white matter]. Neurochirurgie 57:151–155. CrossRef Medline

Price CJ, Crinion J (2005) The latest on functional imaging studies of aphasic stroke. Curr Opin Neurol 18:429–434. CrossRef Medline

Price CJ, Green DW, von Studnitz R (1999) A functional imaging study of translation and language switching. Brain 122:2221–2235. CrossRef Medline

Raboyeau G, Marcotte K, Adrover-Roig D, Ansaldo AI (2010) Brain activation and lexical learning: the impact of learning phase and word type. Neuroimage 49:2850–2861. CrossRef Medline

Ranganath C, Heller AS, Wilding EL (2007) Dissociable correlates of two classes of retrieval processing in prefrontal cortex. Neuroimage 35:1663–1673. CrossRef Medline

Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TE (2008) The evolution of the arcuate fasciculus revealed with comparative DTI. Nat Neurosci 11:426–428. CrossRef Medline

Sakai KL, Nauchi A, Tatsuno Y, Hirano K, Muraishi Y, Kimura M, Bostwick M, Yusa N (2009) Distinct roles of left inferior frontal regions that explain individual differences in second language acquisition. Hum Brain Mapp 30:2440–2452. CrossRef Medline

Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C (2008) Ventral and dorsal pathways for language. Proc Natl Acad Sci U S A 105:18035–18040. CrossRef Medline

Schlegel A, Rudelson JJ, Tse P (2012) White matter structure changes as adults learn a second language. J Cogn Neurosci 24:1664–1670. CrossRef

Scholz J, Klein MC, Behrens TE, Johansen-Berg H (2009) Training induces changes in white-matter architecture. Nat Neurosci 12:1370–1371. CrossRef Medline

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 [Suppl 1]:S208–S219. Medline

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TE (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31:1487–1505. CrossRef Medline

Takeuchi H, Sekiguchi A, Taki Y, Yokoyama S, Yomogida Y, Komuro N, Yamanouchi T, Suzuki S, Kawashima R (2010) Training of working memory impacts structural connectivity. J Neurosci 30:3297–3303. CrossRef Medline

Taubert M, Draganski B, Anwander A, Müller K, Horstmann A, Villringer A, Ragert P (2010) Dynamic properties of human brain structure: learning-related changes in cortical areas and associated fiber connections. J Neurosci 30:11670–11677. CrossRef Medline

Tettamanti M, Alkadhi H, Moro A, Perani D, Kollias S, Weniger D (2002) Neural correlates for the acquisition of natural language syntax. Neuroimage 17:700–709. CrossRef Medline

Tettamanti M, Moro A, Messa C, Moresco RM, Rizzo G, Carpinelli A, Matarrese M, Fazio F, Perani D (2005) Basal ganglia and language: phonology modulates dopaminergic release. Neuroreport 16:397–401. CrossRef Medline

Thierry G, Wu YJ (2007) Brain potentials reveal unconscious translation during foreign-language comprehension. Proc Natl Acad Sci U S A 104:12530–12535. CrossRef Medline

Thompson PM, Cannon TD, Narr KL, van Erp T, Poutanen VP, Huttunen M, Lönnqvist J, Standertskjöld-Nordenstam CG, Kaprio J, Khaledy M, Dail R, Zoumalan CI, Toga AW (2001) Genetic influences on brain structure. Nat Neurosci 4:1253–1258. CrossRef Medline

van Ettinger-Veenstra HM, Ragnehed M, Hällgren M, Karlsson T, Landtblom AM, Lundberg P, Engström M (2010) Right-hemispheric brain activation correlates to language performance. Neuroimage 49:3481–3488. CrossRef Medline

Van Ettinger-Veenstra H, Ragnehed M, McAllister A, Lundberg P, Engström M (2012) Right-hemispheric cortical contributions to language ability in healthy adults. Brain Lang 120:395–400. CrossRef Medline

Veroude K, Norris DG, Shumskaya E, Gullberg M, Indefrey P (2010) Functional connectivity between brain regions involved in learning words of a new language. Brain Lang 113:21–27. CrossRef Medline

Videsott G, Herrnberger B, Hoenig K, Schilly E, Grothe J, Wiater W, Spitzer M, Kiefer M (2010) Speaking in multiple languages: neural correlates of language proficiency in multilingual word production. Brain Lang 113:103–112. CrossRef Medline

Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N (2011) What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. Neuroimage 54:577–593. CrossRef Medline

Vingerhoets G, Van Borsel J, Tesink C, van den Noort M, Deblaere K, Seurinck R, Vandemaele P, Achten E (2003) Multilingualism: an fMRI study. Neuroimage 20:2181–2196. CrossRef Medline

Wan X, Takano D, Asamizuya T, Suzuki C, Ueno K, Cheng K, Ito T, Tanaka

K (2012) Developing intuition: neural correlates of cognitive-skill learning in caudate nucleus. J Neurosci 32:17492–17501. Medline

Wong FC, Chandrasekaran B, Garibaldi K, Wong PC (2011) White matter anisotropy in the ventral language pathway predicts sound-to-word learning success. J Neurosci 31:8780–8785. CrossRef Medline

Yeatman JD, Dougherty RF, Rykhlevskaia E, Sherbondy AJ, Deutsch GK,

Wandell BA, Ben-Shachar M (2011) Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. J Cogn Neurosci 23:3304–3317. CrossRef Medline

Zatorre RJ, Fields RD, Johansen-Berg H (2012) Plasticity in gray and white: neuroimaging changes in brain structure during learning. Nat Neurosci 15:528–536. CrossRef Medline