Brain Struct Funct (2016) 221:577–589
DOI 10.1007/s00429-014-0926-2

ORIGINAL ARTICLE

White matter pathway supporting phonological encoding in speech production: a multi-modal imaging study of brain damage patients

Zaizhu Han · Yujun Ma · Gaolang Gong · Ruiwang Huang · Luping Song · Yanchao Bi

Received: 30 March 2014 / Accepted: 17 October 2014 / Published online: 31 October 2014

Abstract In speech production, an important step before motor programming is the retrieval and encoding of the phonological elements of target words. It has been proposed that phonological encoding is supported by multiple regions in the left frontal, temporal and parietal regions and their underlying white matter, especially the left arcuate fasciculus (AF) or superior longitudinal fasciculus (SLF). It is unclear, however, whether the effects of AF/SLF are indeed related to phonological encoding for output and whether there are other white matter tracts that also contribute to this process. We comprehensively investigated the anatomical connectivity supporting phonological encoding in production by studying the relationship between the integrity of all major white matter tracts across the entire brain and phonological encoding deficits in a group of 69 patients with brain damage. The integrity of each white matter tract was measured both by the percentage of damaged voxels (structural imaging) and the mean fractional anisotropy value (diffusion tensor imaging). The phonological encoding deficits were assessed by various measures in two oral production tasks that involve phonological encoding: the percentage of nonword (phonological) errors in oral picture naming and the accuracy of word reading aloud with word comprehension ability regressed out. We found that the integrity of the left SLF in both the structural and diffusion tensor imaging measures consistently predicted the severity of phonological encoding impairment in the two phonological production tasks. Such effects of the left SLF on phonological production remained significant when a range of potential confounding factors were considered through partial correlation, including total lesion volume, demographic factors, lesions on phonological-relevant grey matter regions, or effects originating from the phonological perception or semantic processes. Our results therefore conclusively demonstrate the central role of the left SLF in phonological encoding in speech production.

Keywords Phonological encoding · Left superior longitudinal fasciculus · Diffusion tensor imaging · Patients · Connectome

Z. Han, Y. Ma are co-first authors.

Z. Han (✉) · Y. Ma · G. Gong · Y. Bi
State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China
e-mail: zzhhan@bnu.edu.cn

Z. Han · G. Gong · Y. Bi
Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing 100875, China

Y. Ma
Department of Psychology, Henan University, Kaifeng 475000, China

R. Huang
Center for the Study of Applied Psychology, Key Laboratory of Mental Health and Cognitive Science of Guangdong Province, School of Psychology, South China Normal University, Guangzhou 510631, China

L. Song (✉)
Department of Neuro-rehabilitation, Rehabilitation College of Capital Medical University and China Rehabilitation Research Center, Beijing 100068, China
e-mail: songluping882002@yahoo.com.cn

Introduction

To produce a spoken word, upon the selection of the target lexical item from the mental lexicon, its phonological

elements (e.g., phonemes) need to be identified and ordered before motor program planning and execution (e.g., Dell 1986; Levelt 1999). A failure in the phonological encoding commonly results in producing nonword items that are often phonologically related to the target in a wide range of oral production tasks, such as spontaneous speech, oral naming, word reading aloud, and repetition (e.g., saying "taple" for "table").

Neuropsychological and functional neuroimaging studies have commonly associated phonological encoding in production to a set of cortical regions in the left frontal, temporal and parietal cortices, including the supplementary motor area, the left anterior insula, the left posterior superior and middle temporal gyri (Wernicke's area), and the left posterior superior temporal gyrus (Anderson et al. 1999; Buchsbaum et al. 2001; Indefrey and Levelt 2004). In an influential dual-stream model of speech processing (Hickok and Poeppel 2007), these regions constitute the dorsal pathway that maps acoustic speech signals to frontal lobe articulatory networks, which underlie phonological encoding processing for output. Schwartz et al. (2012) used a voxel-based lesion-symptom mapping (VLSM) approach to examine the lesion patterns associated with the percentage of nonword (phonological) errors in oral picture naming, which was assumed to arise from the phonological encoding stage for output. They observed that nonword (phonological) errors in picture naming were associated with lesions in the anterior part of the dorsal stream (i.e., premotor, pre and postcentral gyrus and supermarginal gyrus), and not the auditory-related posterior temporal and temporo-parietal cortices, which was argued to be consistent with the motor-control models of speech production (Tremblay et al. 2003; see also Hickok and Poeppel 2004, 2007).

These phonological-encoding relevant cortical regions are generally associated with the classical disconnection syndrome (Compston 2006; Geschwind 1965; Wernicke 1970). They tend to be situated on the white matter tract (arcuate fasciculus, AF, or superior longitudinal fasciculus, SLF[1]) that connects Wernicke's and Broca's areas, whose disruption classically results in conduction aphasia, a syndrome with oral repetition difficulty as the signature symptom. However, it is premature to assume that the AF/SLF is the tract supporting phonological encoding for output. Although conduction aphasia may result from phonological encoding deficits, the syndrome has high heterogeneity with potentially different types of cognitive impairment (see Dick and Tremblay 2012 for a review). Indeed, recent studies with more precise in vivo diffusion tensor imaging (DTI) have shown conflicting results regarding whether AF/SLF damage is associated with conduction aphasia symptoms, especially repetition difficulty (Mendez and Benson 1985; Saur et al. 2008; Selnes et al. 2002; Shuren et al. 1995; see review in Dick and Tremblay 2012). Similarly, other studies reporting phonological-related white matter tracts have often used tasks that involve not only phonological encoding but also other cognitive processes such as speech perception (repetition; Breier et al. 2008; Fridriksson et al. 2010; Rolheiser et al. 2011; Saur et al. 2008) or lexical retrieval (e.g., speech fluency, Marchina et al. 2011), and therefore, it is unclear whether these tracts are responsible for phonological encoding or for other language processes. The behavioral measure used in Schwartz et al. (2012) was likely immune to this concern, and they showed that the effective regions had decent overlap with AF/SLF in the Johns Hopkins University Probabilistic Atlas (JHU WM template, http://cmrm.med.jhmi.edu) (29 % of the entire SLF tract). However, without DTI data, the effects of white matter tracts were not examined directly. Mandonnet et al. (2007) used intraoperative electrical stimulation and found that stimulating the AF, and not the inferior fronto-occipital fasciculus/inferior longitudinal fasciculus, induced phonological paraphasia in oral picture naming in 7 of 12 patients. Although highly informative, the results require further scrutiny because (1) some phonological errors may be real word errors (e.g., saying "house" for the picture of "horse") and could stem from failures in the lexical selection stage (e.g., Gagnon et al. 1997); (2) patients with glioma (in the case of a slow growing tumor) may undergo structural/functional reorganization (Goebell et al. 2006; Lazar et al. 2006); and (3) highly limited tracts of prior interest were targeted without the opportunity to reveal the potential effects of other tracts. Therefore, studies combining more comprehensive examination of the white matter tracts and measures specifically reflecting phonological encoding are needed.

We investigated the anatomical connectivity supporting phonological encoding in speech production by studying the relationship between the integrity of all major white matter tracts and phonological encoding deficits in a group of 69 patients with brain damage. To assess white matter integrity, we measured the percentage of lesion voxels (structural imaging) and the mean fractional anisotropy (FA) value (diffusion tensor imaging) of each major tract. For the behavioral measures, we employed two tasks that involved the phonological encoding stage: oral picture naming and word reading aloud. For picture naming, we followed Schwartz et al. (2012) and used the percentage of

[1] Note that although AF and SLF have been traditionally viewed as synonyms depicting the same tract (see Dick and Tremblay 2012 for a review), there are recent debates about their distinctions (Makris et al. 2005; Petrides and Pandya 2009; Schmahmann et al. 2007). We for now do not distinguish them to cover studies about these tracts more comprehensively and defer their potential differences to the discussion.

nonword (phonological) errors. For word reading aloud, given the dual-route architecture and the interactive nature of the semantic and nonsemantic routes in reading (e.g., Bi et al. 2007; Hillis and Caramazza 1995), reading errors may originate from complex sources, including, but not restricted to, visual word perception, orthographical lexical access, semantic processing, phonological lexical access, phonological encoding, and orthography-phonology conversion. Therefore, from the word reading accuracy, we regressed out the accuracy of the word associative matching task to partial out the effect originating from word comprehension stages (visual word perception, orthographic and semantic processing). The resulting residual measure would reflect integrity of the remaining systems, including post-lexical phonological encoding. Potential confounding factors, such as whole-brain lesion volume, demographic variables (age, gender, years of education) and the effect of relevant gray matter regions (those associated with nonword errors in oral picture naming), were examined in results validation analyses. Two control tasks (auditory lexical decision, picture associative matching) were also included to control for the effects of the phonological perception and comprehension stages. Our rationale is that if a tract is necessary for phonological encoding in speech production, its integrity measures (lesion percentage, mean FA value) should significantly correlate with the phonological encoding indices in oral picture naming and word reading aloud across patients, even when controlling for potentially confounding variables.

Materials and methods

Participants

Seventy-nine patients with brain damage (65 males) participated in the study. They were all native Chinese speakers and were recruited from the China Rehabilitation Research Center based on the following inclusion criteria: willingness to participate; no previous brain injury or other neurological/psychiatric disease such as alcohol abuse or severe depression; at least one-month post-onset [disease duration: mean = 7 months; standard deviation (SD) = 12; range 1–86]; ability to follow task instructions; and a relatively clear lesion profile (stroke and traumatic brain injury). We did not constrain our patient selection to those clinically diagnosed with aphasia so that more patients with a greater variety of lesion coverage were included to allow for statistical analyses and potential discoveries beyond the conventional aphasic studies. Analyses were also performed on only the stroke patients. Seven patients were not included in the analyses because of poor signal quality of the DTI images (one for excessive head movement; others for normalization failure). Three other patients were excluded because they had severe apraxia of speech and gave no meaningful response in the production tasks. The mean age of the remaining 69 patients (55 males) was 46 years (SD = 13; range 20–76), and the mean formal education was 13 years (SD = 3; range 2–19). The majority of the patients ($n = 63$) suffered from stroke, and 6 suffered from traumatic brain injury. According to Gao (1993), neuropsychological tests of language processing revealed that 8 patients did not suffer from aphasia, and the remaining patients suffered from motor (21 %), sensory (13 %), anomia (20 %), conduction (5 %), global/mixed (34 %), and subcortical aphasia (8 %). Administration of the Edinburgh Handedness Inventory (Oldfield 1971) revealed that all patients were right-handed. The Chinese version of the Mini-mental state examination (MMSE) (Folstein et al. 1975) was administered (mean = 22; SD = 8; range 3–30). All participants provided written informed consent. This study was approved by the Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University.

Behavioral data

Phonological encoding ability in oral production was assessed by two common tasks that both involve this cognitive component: oral picture naming and word reading aloud. These two tasks included pictures or names of a variety of common objects (see below for details) and were considered in parallel for convergence. For oral picture naming, the percentage of nonword errors was used as an estimate of phonological encoding impairment; for word reading aloud, the response accuracy after regressing out word associative matching (i.e., word comprehension) was used.

The experimental tasks were conducted in separate sessions using the DMDX program (Forster and Forster 2003) on a personal computer. For each task, the item presentation order was randomized and was constant across subjects. There was a 60-second response deadline. Participants were tested individually in a quiet room. Each session lasted less than 2 hours, and the tests were paused for rest upon patient request. The responses in both tasks were taped using two digital recorders and were transcribed for scoring. For all tasks, the response for each item was scored 1 if correct or 0 if wrong or if no response given within the deadline. The erroneous responses in oral picture naming were further classified into nonword errors and other types of errors (see below).

Oral picture naming

Color photographs of 100 common objects were used, including 20 items from each of five categories: animals (e.g., dog), tools (e.g., hammer), common artifacts (e.g., chair), fruits and vegetables (e.g., apple), and large non-manipulable objects (e.g., house). Participants were instructed to name each object. The first complete response was scored. To tap into the post-lexical phonological encoding stage, we followed the measures used by Schwartz et al. (2012), focusing on non-word errors and including non-word errors that were either phonologically related or unrelated to the target. Real word phonological error responses were not included because they may originate from lexical selection failures (Caramazza 1997). Following Schwartz et al. (2012), we calculated for each patient the proportion of nonword errors in total trials ($n = 100$) and then normalized the value by taking the square root. We further negated the obtained value so that higher scores reflected better phonological encoding abilities. We also performed analyses on only the non-word phonological errors, defined as being phonologically related to the target, i.e., sharing at least one phoneme in the same position (e.g., ban ma→guan ma) or two phonemes in any position (e.g., xiang ce→ce xiang) excluding schwa. The result patterns were fully consistent.

Word reading aloud

The written names of the objects used in the picture naming task were presented to the patients to read aloud. As described in the Introduction, we estimated the effects arising from the post-lexical encoding stage by regressing out the performance on a visual word recognition task: word associative matching. The matching task had the same format as the Pyramids and Palm Trees Test (Howard and Patterson 1992), with each trial containing the written names of three objects on a touch screen. Participants judged which of the two bottom names (e.g., orange, onion) was semantically closer to the top one (e.g., banana) by pressing the touch screen. There were 50 trials with ten items from each of the 5 semantic categories in the oral picture naming task. The three items in each trial were always from the same category. In this task, the participants needed to recognize the visual word, activating the orthographic and semantic representation (Weekes et al. 2006). Phonological encoding ability was estimated by calculating the residual of word reading accuracy against the word comprehension accuracy. Specifically, we first established a regression model using patient data, with the dependent variable being raw word reading accuracy and the predictor being raw word recognition accuracy. Then, for each patient, a predicted accuracy score of word reading was acquired by introducing his or her raw word recognition accuracy into the model. This patient's word reading residual value equaled his/her raw accuracy minus predicted accuracy. The residual values across all patients summed to zero.

Control tasks

To assess the (relative) specificity of potential phonological-output related tracts, we included two control tasks. One required phonological perception but not necessarily phonological output processing: auditory lexical decision; the other required semantic processing with no or minimal phonological processing: picture associative matching. The auditory lexical decision task had ten items. For each item, the patients were presented a Chinese word or nonword auditorily via the earphone and needed to decide whether the stimulus was a real word in Chinese by pressing the "YES" or "NO" button. The picture associative matching task was identical to the above word associative matching task except that the written names of objects in the word task were replaced by the photographs of the objects. Specifically, no verbal processing was needed in the completion of this task.

Imaging data

Acquisition

Patients were scanned at the China Rehabilitation Research Center with a 1.5T GE SIGNA EXCITE scanner. Three types of images were collected: (1) high resolution three-dimensional T1-weighted images (3D images), (2) T2-weighted fluid attenuated inversion recovery (FLAIR) images, and (3) diffusion-weighted images (DWI). The T1-weighted 3D images were acquired along the sagittal plane with the following parameters: matrix size = 512 × 512, voxel size = 0.49 × 0.49 × 0.70 mm^3, repetition time (TR) = 12.26 ms, echo time (TE) = 4.2 ms, inversion time (TI) = 400 ms, field of view (FOV) = 250 × 250 mm^2, flip angle = 15°, slice number = 248 slices. The T2-weighted FLAIR images were obtained on the axial plane with the following parameters: matrix size = 512 × 512, voxel size = 0.49 × 0.49 × 5 mm^3, TR = 8,002 ms, TE = 127.57 ms, TI = 2,000 ms, FOV = 250 × 250 mm^2, flip angle = 90°, slice number = 28 slices. The DWIs were acquired in two separate sequences with different diffusion weighting direction sets. The first sequence was acquired using the following parameters: 15 diffusion weighting directions, matrix size = 128 × 128, voxel size = 1.95 × 1.95 × 2.6 mm^3, TR = 13,000 ms, TE = 69.3 ms, TI = 0 s, FOV = 250 × 250 mm^2, flip angle = 90°, slice number = 53

slices, slice thickness = 2.6 mm without gap. The other sequence was acquired using the same parameters, except that it included 17 different directions. The first two volumes were b0 volumes, and the *b* value of other volumes was 1,000 s/mm^2 in each sequence. All the sequences except for the T2-weighted FLAIR sequence were scanned twice in the same session to improve the signal-to-noise ratio.

Preprocessing

For the T1-weighted 3D images, we first co-registered each of the two datasets on the native space and averaged them using SPM5 (http://www.fil.ion.ucl.ac.uk/spm/software/spm5). The T2-weighted FLAIR images were then coregistered and resliced to the averaged 3D images with SPM5. Two trained researchers manually drew each patient's lesion contour on the averaged 3D images slice by slice, visually referring to the FLAIR images. The lesion-drawing procedure was supervised by an experienced radiologist. Each patient's structural images were resliced into $1 \times 1 \times 1$ mm^3 voxels and then registered into Talairach space via BrainVoyager QX v2.0 (www.brainvoyager.com). We used the ANTs software package (Advanced Normalization Tools, http://www.picsl.upenn.edu/ANTS/) to extract the affine transformation matrix between the native and Talairach spaces and to register the lesion description in Talairach space. Additionally, using the ANTs, the lesion description was transformed into the Montreal Neurological Institute (MNI) space.

Regarding the DWI data, for each patient, we first merged each of the 15 directions and 17 directions paired sequences into one single 4D nifti-1 format file and merged the diffusion weighted gradient tables of the two sequences. We then executed the following steps using a pipeline tool, PANDA (Cui et al. 2013), which was based on FSL4.1.7 (http://www.fmrib.ox.ac.uk/fsl/index.html/): BET: skull removal; Eddycorrect: correction of eddy current distortion; DTIFIT: build diffusion tensor model. After obtaining the FA maps of each patient, we registered them to the FMRIB FA template in MNI space using ANTs (version 1.9). The normalization included two steps: linear affine and nonlinear transform registration. In linear affine transform, the *WarpImageMultiTransform* program was executed, and one affine transform txt file for each participant was produced. In nonlinear transform, a shell script "buildtemplate" was used. We thus obtained a normalized FA map of each patient.

The index of white matter tract integrity

We focused on the major tracts in a widely used human brain white-matter template, the "JHU white-matter tractography atlas" (25 % probability map) from FSL (http://www.fmrib.ox.ac.uk/fsl/data/atlas-descriptions.html#wm), which contains 20 main white matter bundles (see Table 1 for details). Our analyses considered two indices reflecting the severity of brain damage: lesion percentage and FA value. The lesion percentage index measures the percentage of voxels that have physical damage, wherein a given voxel is scored as intact or damaged. It is a classical indicator of lesion severity (Bates et al. 2003). FA value is one of the most widely used variables to evaluate the integrity of white matter tracts in a given voxel (Rolheiser et al. 2011; Wilson et al. 2011), reflecting fiber density, axonal diameter, and myelination in white matter (Basser and Pierpaoli 1996). In our dataset, each voxel in each patient had a lesion value (categorical variable: lesion or intact) from the T1 lesion map and a FA value (continuous variable) from the normalized FA map.

Brain-behavior mapping analysis

To identify the major white matter pathways responsible for phonological encoding, we examined the relationship between the integrity of the major tracts (measured by lesion percentage and FA value) and phonological encoding ability.

First, for each behavior index, we correlated the lesion percentage and the mean FA value of each tract and phonological encoding index, with whole-brain total lesion volume as a covariate. For the lesion percentage measure, three tracts were lesioned in fewer than five patients and were excluded from the lesion analysis. For each of the remaining 17 tracts, the lesion percentage (number of voxels with lesion divided by total number of voxels in the tract) was correlated with the patients' phonological encoding scores in oral picture naming and word reading, partialing out total lesion volume. The results were adjusted for the 17 tracts using the Bonferroni correction method ($p < 0.0006$, corrected $p < 0.01$). For the FA measure, the mean FA value of each of the 20 tracts in the template was obtained by averaging the FA values of all voxels in the tract. The mean FA value was correlated with the behavioral scores for each task across patients, controlling for total lesion volume. The Bonferroni correction method ($p < 0.0005$, corrected $p < 0.01$) was implemented (20 tracts).

Next, we considered a set of potentially confounding variables in addition to total lesion volume—demographic and effects of relevant gray matter—by performing the following analyses (see Table 2): (1) partial correlation between phonological encoding scores and the tract lesion percentages and mean FA values, with patients' age, gender, education level (total years of formal education period), and total lesion volume as covariates; (2) partial

Table 1 Correlation coefficients between lesion percentages and mean FA values of each tract and the two phonological encoding measures across 69 patients, with total lesion volume as covariate

Tract		Total volume (mm^3)	# of patients with lesion	Lesion-behavior correlation		FA-behavior correlation	
				Oral picture naming index	Oral word reading index	Oral picture naming index	Oral word reading index
1	Left anterior thalamic radiation (ATR)	8,128	41	−0.28	−0.19	0.32	0.20
2	Right anterior thalamic radiation (ATR)	7,576	32	0.29	0.33	−0.46*	−0.37
3	Left corticospinal tract (CST)	5,464	40	−0.29	−0.28	0.50*	0.37
4	Right corticospinal tract (CST)	4,760	26	0.41*	0.37	−0.48*	−0.39
5	Left cingulum gyrus (CCG)	1,552	4	–	–	−0.06	−0.04
6	Right cingulum gyrus (CCG)	608	7	0.16	0.22	−0.25	−0.26
7	Left cingulum hippocampus (CH)	248	0	–	–	−0.04	−0.05
8	Right cingulum hippocampus (CH)	544	2	–	–	−0.24	−0.20
9	Forceps major (F_MA)	5,744	17	0.24	0.28	−0.20	−0.21
10	Forceps minor (F_MI)	19,712	41	0.07	0.07	−0.19	−0.13
11	Left inferior fronto-occipital fasciculus (IFOF)	5,048	40	−0.31	−0.19	0.38	0.32
12	Right inferior fronto-occipital fasciculus (IFOF)	6,304	34	0.37	0.40	−0.43*	−0.44*
13	Left Inferior longitudinal fasciculus (ILF)	5,400	24	0	−0.29	0.17	0.28
14	Right Inferior longitudinal fasciculus (ILF)	3,152	9	0.25	0.29	−0.36	−0.39
15	Left superior longitudinal fasciculus (SLF)	9,472	42	−0.52*	−0.57*	0.52*	0.55*
16	Right superior longitudinal fasciculus (SLF)	7,456	19	0.28	0.40	−0.44*	−0.44*
17	Left uncinate fasciculus (UF)	744	21	−0.26	−0.14	0.30	0.29
18	Right uncinate fasciculus (UF)	448	15	0.29	0.28	−0.40	−0.38
19	Left superior longitudinal fasciculus (temporal part) (SLFT)	96	7	−0.10	−0.25	0.21	0.34
20	Right superior longitudinal fasciculus (temporal part) (SLFT)	72	6	0.20	0.27	−0.28	−0.43*

Oral picture naming index: negated square root for the percentage of nonword errors in all items; Oral word reading index: residual of word reading accuracy after regressing out the word associative matching accuracy

Bonferroni corrected: * $p < 0.01$; "–"could not carry out this analysis due to few patients

correlation between phonological encoding scores and the tract lesion percentages and mean FA values, with the lesion percentage in the gray matter mask related to phonological encoding and total lesion volume as covariates.

The phonological-encoding-relevant gray matter regions were acquired using the following method. First, we followed the Schwartz et al. (2012) study and conducted a VLSM (Bates et al. 2003) analysis for the oral picture

Table 2 Partial correlation coefficients between the integrity of the left SLF (lesion percentages and mean FA values) and behavioral measures, controlling for potential confounding factors

Control aspect	Control variable	Analysis type	Oral picture naming index	Oral word reading index
Demographic factor	Age, education, gender, total lesion volume	Lesion analysis	−0.50***	−0.56***
		FA analysis	0.50***	0.54***
Influence of gray matter	Lesion volume in grey matter regions of VLSM significant results, total lesion volume	Lesion analysis	−0.33**	−0.47***
		FA analysis	0.36**	0.46**
Variant of brain damage	Only 63 right-handed stroke patients: total lesion volume	Lesion analysis	−0.50***	−0.57***
		FA analysis	0.50***	0.55***
Control tasks	Auditory lexical decision score, total lesion volume	Lesion analysis	−0.49***	−0.56***
		FA analysis	0.49***	0.54***
	Picture associative matching score, total lesion volume	Lesion analysis	−0.51***	−0.59***
		FA analysis	0.53***	0.56***

Oral picture naming and oral word reading indices were identical to those in Table 1

**$p < 0.01$; ***$p < 0.001$

naming index from the 69 patients. We used the nonparametric mapping program MRIcroN (Rorden et al. 2007) to obtain a VLSM map of a nonparametric Brunner-Munzel test (Brunner and Munzel, 2000), on which the threshold was set at false discovery rate (FDR) correction $p < 0.01$. The resulting whole brain VLSM map was then overlaid on a gray matter mask (SPM5 template, probability higher than 0.4).

We then examined whether the results were influenced by the inclusion of different types of etiology by analyzing only the stroke patients. Partial correlations were computed between patients' phonological encoding scores and lesion percentages or mean FA values, with total lesion volume value as a covariate for only the 63 patients with stroke.

Furthermore, we tested whether the phonological related tracts were (relatively) specific to phonological production by considering two other processes (phonological perception and semantic processing). We examined the association between phonological-encoding-related tract integrity (lesion percentage and FA value) and performance on a phonological perception task (auditory lexical decision). For each tract, the auditory lexical decision scores were correlated with the lesion percentages or mean FA values across patients, with total lesion volume value as a covariate. We then examined whether there was phonological encoding effect over and above any potential effects of this control task using a partial correlation analysis between phonological indices and lesion percentages or FA values, with total lesion volume and the control task performance as covariates. The same analysis procedure was also applied to the picture associative matching semantic control task.

Finally, although separate analyses were performed to more clearly address specific questions and to avoid a collinearity problem introduced by having too many covariates, we nonetheless reported the results with all of the above control variables in one analysis for convergence. For the tracts that had been observed to have significant effects in the above analyses, we correlated their integrity values with the phonological encoding scores in the 63 stroke patients, partialling out age, gender, education level, lesion volume in gray matter regions of VLSM significant results, auditory lexical decision score, picture associative matching score, and total lesion volume.

Results

Behavioral performance

The response accuracy of the 69 patients in the oral picture naming task was 75 ± 26 % (mean ± SD). The percentage of nonword (phonological) errors in this task was 6 ± 12 %, and the negated normalized phonological encoding score in this task was −15 ± 20 %. The accuracy of the word reading task was 86 ± 19 %, and that of the word associative matching task was 89 ± 9 %. The phonological encoding score in this task (i.e., the residual of word reading accuracy, regressing out the accuracy of the word associative matching task) was 0 ± 16.71 %. The two phonological-encoding scores were positively correlated ($r = 0.45$, $p < 0.0001$). The accuracies for the auditory lexical decision and picture associative matching control tasks were 90 ± 11 % and 87 ± 11 %, respectively.

Brain-behavior mapping

The lesion distribution of the 69 patients is presented in Fig. 1. All white matter tracts except for three [left

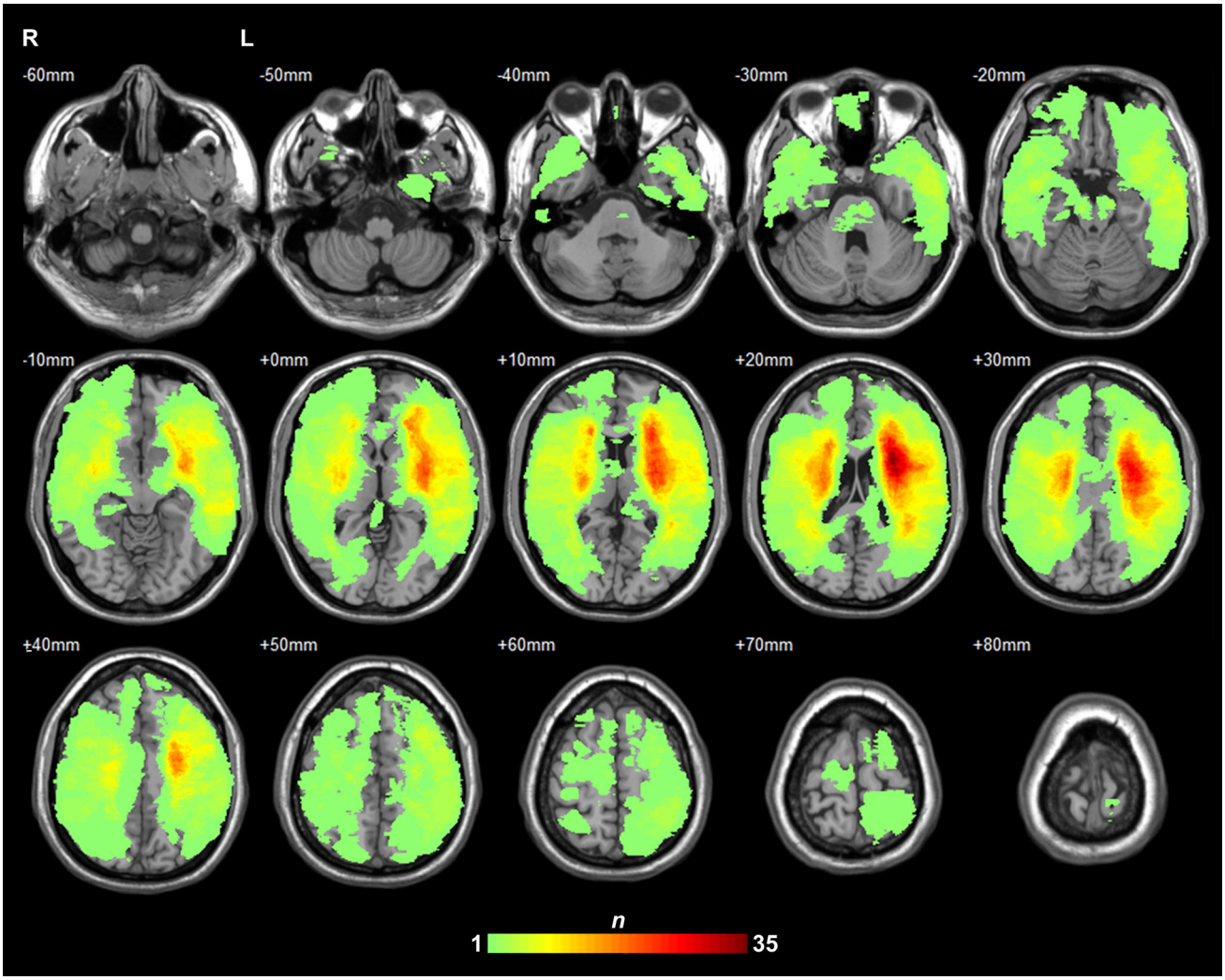

Fig. 1 Lesion overlap map of the 69 patients (the *n* value on each voxel denotes the number of patients with a lesion on it)

cingulum gyrus, left and right cingulum (hippocampus)] were lesioned in more than five patients. Five tracts (left anterior thalamic radiation, left corticospinal tract, left inferior fronto-occipital fasciculus, forceps minor, and left SLF) were damaged (having >1 voxel with lesions) in more than 50 % of the patients, and five additional tracts (left inferior longitudinal fasciculus, left uncinate fasciculus, right anterior thalamic radiation, right corticospinal tract, and right inferior fronto-occipital fasciculus) were lesioned in 30–50 % of the patients (Table 1).

Lesion-behavior correlation

As observed in Table 1 and Fig. 2, the lesion percentages of the left SLF were significantly negatively correlated with phonological encoding scores in both the oral picture naming and word reading tasks, with total lesion volume as a covariate (Bonferroni corrected $ps < 0.01$). The lesion percentages of the right corticospinal tract were significantly positively correlated with phonological encoding scores in the oral picture naming (Bonferroni corrected $p < 0.01$), which was opposite the expected direction, possibly due to artifacts (larger right hemisphere lesion correlated with smaller left hemisphere lesion), and was not considered further.

FA-behavior correlation

The mean FA values of the left SLF significantly positively correlated with phonological encoding scores in both the oral picture naming and word reading tasks, controlling for total lesion volume (Bonferroni corrected $ps < 0.01$; Table 1; Fig. 2). Additionally, the mean FA values of the left corticospinal tract correlated with the phonological encoding score from the oral picture naming, partialling out total lesion volume (Bonferroni corrected

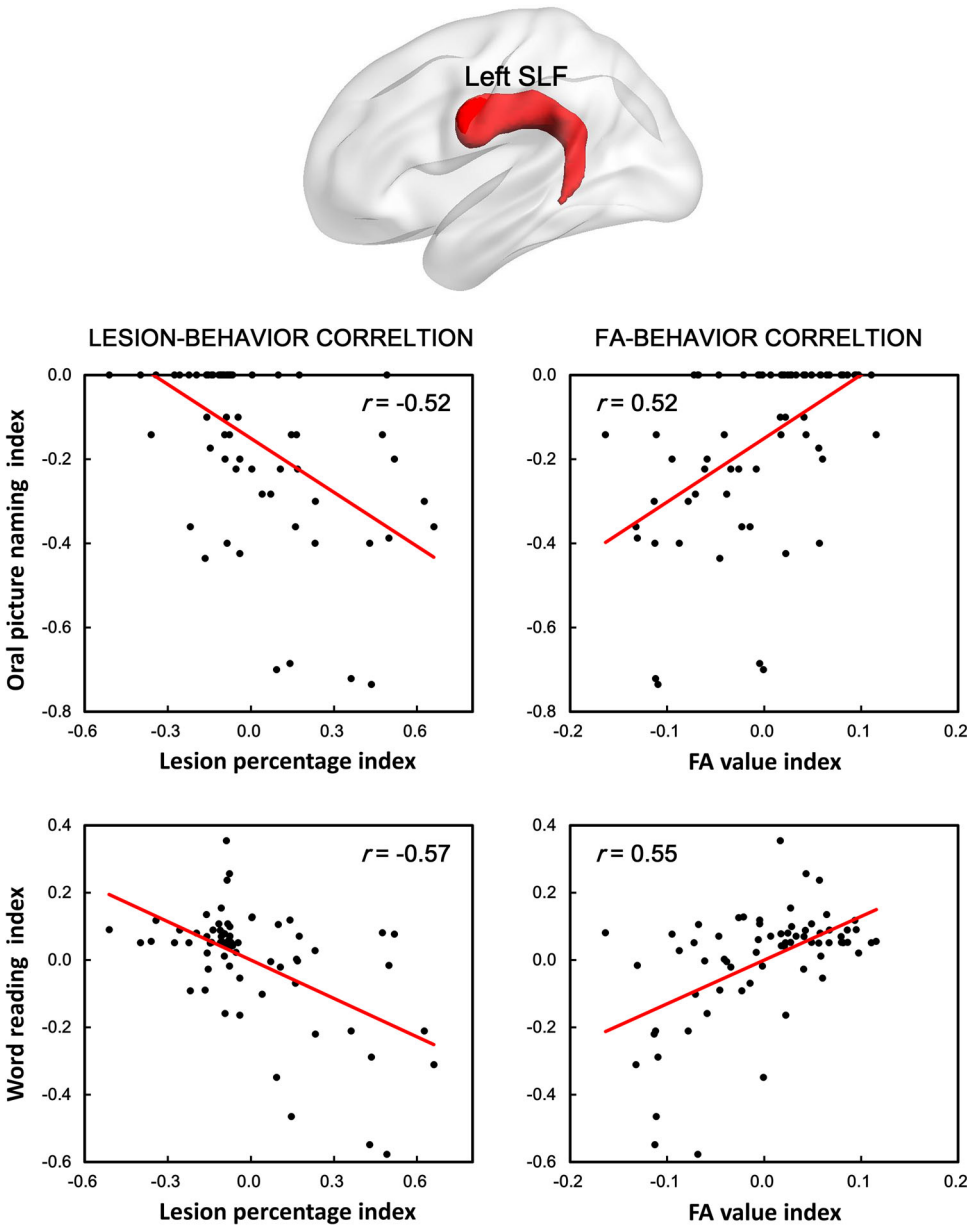

Fig. 2 Correlation between white matter integrity (lesion percentage and mean FA value) of the left superior longitudinal fasciculus (SLF) and phonological encoding measures. Oral picture naming index: negated square root for the percentage of nonword errors in all items; oral word reading index: residual of word reading accuracy after regressing out the word associative matching accuracy; lesion percentage and FA value indices: residuals of lesion percentage and mean FA value after regressing out total lesion volume, respectively

$p < 0.01$). We also identified five right tracts [anterior thalamic radiation, corticospinal tract, inferior fronto-occipital fasciculus, SLF, superior longitudinal fasciculus (temporalpart)] whose mean FA values were negatively correlated with scores for the oral picture naming and/or word reading task, controlling for total lesion volume (Bonferroni corrected $ps < 0.01$); these results were again opposite the expected direction and were not considered further.

Given that the two phonological encoding measures revealed significant effects of the left SLF across the lesion- and FA-behavior correlation analyses, we focused on this tract in the following validation analyses.

Controlling for effects of potential confounding variables

The ability of the left SLF to predict phonological encoding score remained significant even when the effects of a series of potential confounding variables were considered (see Table 2). First, after age, gender and education and total lesion volume were treated as covariates, the phonological encoding scores remained significantly correlated with lesion percentages (oral picture naming: partial $r = -0.50$, $p < 3 \times 10^{-5}$; word reading: partial $r = -0.56$, $p < 2 \times 10^{-6}$) and mean FA values of the left SLF (oral picture naming: partial $r = 0.50$, $p < 3 \times 10^{-5}$; word reading: partial $r = 0.54$, $p < 4 \times 10^{-6}$). Second, using the

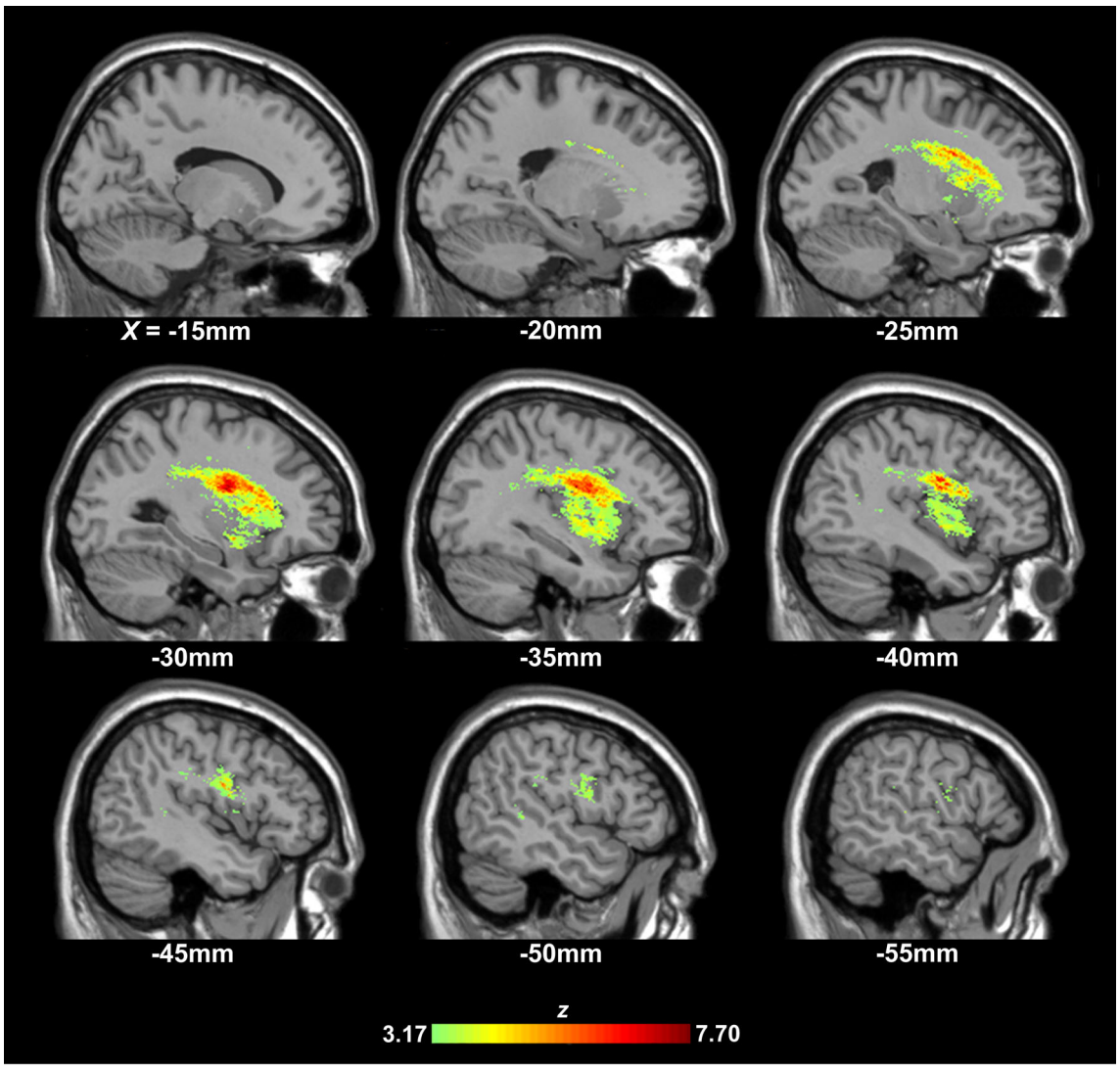

Fig. 3 VLSM statistical map. Results of the nonparametric Brunner-Munzel test of non-word errors in picture naming between patients with and without a lesion on each voxel. The map is thresholded at *FDR* $q = 0.01$ ($z = 3.17$)

VLSM approach, we observed that the lesions in the following left hemispheric grey-matter regions were associated with phonological encoding scores: premotor cortex, pre- and postcentral gyri and supramarginal gyrus (Fig. 3). These results are consistent with findings reported by Schwartz et al. (2012). When controlling for the lesion volume in these gray matter regions as well as total lesion volume, the integrity of the left SLF still tended to correlate with phonological encoding scores for oral picture naming (lesion percentage: partial $r = -0.33$, $p < 0.007$; mean FA: partial $r = 0.36$, $p < 0.004$) and oral reading (lesion percentage: partial $r = -0.47$, $p < 6 \times 10^{-5}$; mean FA: partial $r = 0.46$, $p < 2 \times 10^{-4}$). Finally, the effects of SLF held up well when we considered only the 63 right-handed stroke patients, with total lesion volume value as a covariate (lesion percentage: oral picture naming partial $r = -0.50$, $p < 4 \times 10^{-5}$; word reading partial $r = -0.57$, $p < 2 \times 10^{-6}$; mean FA: oral picture naming partial $r = 0.50, p < 4 \times 10^{-5}$; word reading partial $r = 0.55$, $p < 4 \times 10^{-6}$).

Testing the specificity of the phonological tracts

The integrity of the left SLF showed a trend toward correlation with the auditory lexical decision scores when the total lesion volume value was included as a covariate (lesion percentage: partial $r = -0.21$, $p = 0.09$; mean FA: partial $r = 0.24$, $p = 0.05$). More importantly, when scores on this task and total lesion volume values were treated as covariates, the phonological encoding scores still significantly correlated with the integrity of the left SLF (Table 2; lesion percentage: oral picture naming: partial $r = -0.49$, $p < 3 \times 10^{-5}$; word reading: partial $r = -0.56$,

$p < 1 \times 10^{-6}$; mean FA: oral picture naming: partial $r = 0.49$, $p < 3 \times 10^{-5}$; word reading: partial $r = 0.54$, $p < 3 \times 10^{-6}$), suggesting that the role of the left SLF in phonological encoding for production was not to be reduced to its effect in phonological perception.

The integrity of the left SLF did not correlate with the scores for the non-phonological control task, picture associative matching, with total lesion volume as a covariate (lesion percentage: partial $r = -0.13$, $p = 0.28$; mean FA: partial $r = 0.06$, $p = 0.64$). After partialling out scores on this control task and total lesion volume, the phonological encoding scores still significantly correlated with the integrity of the left SLF (Table 2; lesion percentage: oral picture naming: partial $r = -0.51$, $p < 2 \times 10^{-5}$; word reading: partial $r = -0.59$, $p < 2 \times 10^{-7}$; mean FA: oral picture naming: partial $r = 0.53$, $p < 6 \times 10^{-6}$; word reading: partial $r = 0.56$, $p < 7 \times 10^{-7}$), indicating a (relatively) specific effect of the left SLF in phonological encoding.

Finally, after introducing all the above control variables (age, gender, education, lesion volume in grey matter regions of VLSM significant results, auditory lexical decision score, associative matching score and total lesion volume) as covariates together in one analysis, the phonological encoding scores still correlated significantly with the integrity of the left SLF in the 63 stroke patients (lesion percentage: oral picture naming: partial $r = -0.33$, $p < 0.02$; word reading: partial $r = -0.45$, $p < 6 \times 10^{-4}$; mean FA: oral picture naming: partial $r = 0.37$, $p < 0.006$; word reading: partial $r = 0.41$, $p < 0.002$). This finding further confirms the central role of the left SLF in phonological encoding in speech production.

Discussion

In 69 patients with brain damage, we found that the integrity of the left SLF, measured by the percentage of damaged voxels and mean FA values, predicted the severity of phonological encoding impairment for spoken output in two language tasks (oral picture naming, word reading aloud). Because we used convergent measures that more specifically targeted phonological encoding for spoken output, our results show more conclusively that the left SLF plays an indispensable role in phonological encoding in speech production, which was not attributed to its potential effects in other aspects of language processing such as speech perception or semantic processing. The observed effects were not attributed to effects of demographic variables, overall lesion size or relevant grey matter regions in the premotor cortex, pre- and postcentral gyri or supramarginal gyrus, indicating the crucial role of the left SLF in speech production.

Our results are consistent with the classical and contemporary literature (e.g., Duffau et al. 2002; Geschwind 1965; Maldonado et al. 2011; Saur et al. 2008; Vandermosten et al. 2012; Wernicke 1970) that associates the left SLF with phonological processing and deepen our understanding in the following ways. First, most of the evidence associating the left SLF with phonological processing comes from patients with conduction aphasia and/or uses the results of a repetition task. In this case, the left SLF may be responsible for speech perception or phonological working memory rather than being involved in phonological encoding for production. Second, we did not limit the analyses on a specific set of tracts of interest but considered all major tracts across the two hemispheres. Because fewer than 5 patients had a lesion in the left cingulum gyrus, left cingulum hippocampus or right cingulum hippocampus, we were not able to examine these tracts in phonological (output) encoding. The left anterior thalamic radiation, left corticospinal tract, Forceps minor, and left inferior fronto-occipital fasciculus were lesioned in a similar percentage of patients compared to the left SLF, but these tracts still did not exhibit a significant relationship with phonological encoding deficits, suggesting that this process primarily relies on the left SLF.

The use of multiple imaging methodologies allowed us to examine the relative effects of grey matter and white matter tracts more directly. Importantly, using the VLSM approach that treated voxels across the entire brain in the same fashion, we isolated the effects of voxels that affected phonological encoding in the grey matter. We observed that voxels in which lesions were associated with more severe phonological encoding deficits were located in the grey matter of the premotor cortex, pre- and postcentral gyri and supramarginal gyrus. These findings corroborate those reported in Schwartz et al. (2012), suggesting the necessary roles of these cortical regions in phonological encoding. Similar to those authors, we did not observe a strong effect of the posterior superior temporal gyrus, which is consistent with their argument that Wernicke's region may not be part of the phonological encoding network (but see Hickok and Poeppel 2004, 2007). Although in Schwartz et al. (2012) the effect of white matter was evaluated by overlapping analyses between effective voxels and the white matter tract template, this study directly measured phonological encoding ability using DTI imaging, and we were able to more precisely examine the relative anatomical integrity of specific tracts. Our results confirmed their assertion regarding the role of SLF in phonological encoding based on suggestive evidence. Importantly, we were able to further delineate the relative contribution of grey and white matter and found that after controlling for the lesions in the grey matter, the integrity of the SLF was still predictive of phonological encoding

impairment. Specifically, the effect of the SLF was not fully attributed to the grey matter regions, at least to the extent to which our current voxel-based lesion approach can measure. We are not proposing that this tract computes phonology by itself, but rather, it relays information across cortical regions, which compute phonological codes at various stages, and therefore is an indispensable component underlying phonological encoding. There are various hypotheses about the nature of the phonological codes being relayed by the language production dorsal stream. One proposes that the posterior territory (posterior superior temporal gyrus/sulcus, sylvian-parietal-temporal regions) processes words in phonological codes, which is replayed to the frontal regions (Broca's territory) to encode syllabic position of phonemes before syllabic motor code access (e.g., Indefrey and Levelt 2004; Indefrey 2011). An alternative hypothesis postulates that the lexical phonological information is already encoded into a motor code in the temporo-parietal regions and is relayed to the frontal regions in the auditory-motor format (e.g., Hickok and Poeppel 2004). Our findings of the significant association between SLF integrity and errors arising from phonological code encoding indicate that the SLF plays a critical role in phonological code transfer.

To understand how the left SLF/AF participates in phonological encoding, it is crucial to understand the anatomical details of the tract and the type of cortical regions that it connects. In fact, although the left SLF and AF have been viewed as interchangeable in depicting the fiber bundle that connects Wernicke's classical sensory speech area in the posterior superior temporal gyrus and Broca's motor speech area in the inferior frontal regions, controversy has emerged in very recent years with the aid of modern imaging technology (see review in Dick and Tremblay 2012). It has been proposed that the AF is a part of the SLF (e.g., Makris et al. 2005). The current study is constrained by the spatial scale of the John Hopkins University white-matter template and can only be interpreted with the template tract, which was derived from tracking between regions-of-interest in the middle of the posterior limb of the internal capsule and the splenium of the corpus callosum. Further studies investigating the finer-grain structure of the left SLF in phonological encoding are warranted.

Our results shed light on the structural skeleton of language in general. The left SLF/AF has been traditionally viewed as the most prominent language tract; it connects the two classical language regions (Wernicke and Broca), and it is stronger in humans than nonhumans and stronger in adults than children (Friederici 2009). In this study, we showed that it has a central role in phonological production encoding. Although a limited number of items were used ($n = 10$), we observed a marginally significant effect of the SLF and auditory lexical decision performance, suggesting that this tract may also engage in phonological perception processing. Additional aspects of language processing, however, have been shown to be supported by systems beyond this tract; semantics have been shown to rely more strongly on the left inferior fronto-occipital fasciculus, uncinate fasciculus, and anterior thalamic radiation (Han et al. 2013). Morphological and syntactic processing tends to rely on both the dorsal (AF/SLF) and ventral tracts (Rolheiser et al. 2011). Specifically, our results add to the picture that the language system is processed by multiple left hemisphere tracts, and detailed analyses of the subcomponents of language are needed to understand the full scale of the neurobiology of language.

Conclusion

By examining the relationship between the integrity of the major white matter tracts and phonological encoding ability in production tasks across patients, we found that the left SLF is necessary for the functioning of normal phonological encoding for output. The lesion percentage and mean FA value of this tract significantly predicted the severity of phonological encoding deficits, as measured by two tasks involving oral production. The results highlight the role of the left SLF in speech production and the need for further investigation into the finer structures within the left SLF, as well as the integration mechanisms with other tracts for language processing.

Acknowledgments We thank Litao Zhu for help with imaging methodology, Xiaodong Liu for help with statistical analysis and all BNU-CNLab members for data collection and imaging preprocessing, in particular Yangwen Xu, Fangson Liu and Jing Chen. We are also grateful to all research participants. This study was funded by 973 Program (2013CB837300; 2014CB846100), Major Project of National Social Science Foundation (11&ZD186), NSFC (31171073; 31222024; 31221003; 81071149; 81271548; 81371535; 31271115; 81322021), Beijing Natural Science Foundation (Z111 107067311036), NCET (12-0055; 12-0065), Beijing Nova Program (Z121110002512032), Beijing New Medical Discipline Based Group (100270569). Author contributions: ZH, YB and LS designed research; ZH and YM performed research; ZH, YM and GG analyzed data; RH designed imaging sequence; LS recruited and screened patients; ZH and YB wrote the paper.

References

Anderson JM, Gilmore R, Roper S, Crosson B, Bauer RM, Nadeau S et al (1999) Conduction aphasia and the arcuate fasciculus: a reexamination of the Wernicke-Geschwind model. Brain Lang 70(1):1–12

Basser PJ, Pierpaoli C (1996) Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. J Magn Reson B 111(3):209–219

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF (2003) Voxel-based lesion–symptom mapping. Nat Neurosci 6(5):448–450

Bi Y, Han Z, Weekes B, Shu H (2007) The interaction between semantic and the nonsemantic systems in reading: evidence from Chinese. Neuropsychologia 45(12):2660–2673

Breier JI, Hasan KM, Zhang W, Men D, Papanicolaou AC (2008) Language dysfunction after stroke and damage to white matter tracts evaluated using diffusion tensor imaging. Am J Neuroradiol 29(3):483–487

Brunner E, Munzel U (2000) The nonparametric Behrens-Fisher problem: asymptotic theory and a small-sample approximation. Biom J 42(1):17–25

Buchsbaum BR, Hickok G, Humphries C (2001) Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. Cogn Sci 25(5):663–678

Caramazza A (1997) How many levels of processing are there in lexical access? Cogn Neuropsychol 14(1):177–208

Compston A (2006) From the archives. Brain 129(6):1347–1350

Cui Z, Zhong S, Xu P, He Y, Gong G (2013) PANDA: a pipeline toolbox for analyzing brain diffusion images. Front Hum Neurosci 7:42

Dell GS (1986) A spreading-activation theory of retrieval in sentence production. Psychol Rev 93(3):283–321

Dick AS, Tremblay P (2012) Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. Brain 135(12):3529–3550

Duffau H, Capelle L, Sichez N, Denvil D, Lopes M, Sichez JP et al (2002) Intraoperative mapping of the subcortical language pathways using direct stimulations. Brain 125(1):199–214

Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12(3):189–198

Forster KI, Forster JC (2003) DMDX: a Windows display program with millisecond accuracy. Behav Res Method Instrum Comp 35(1):116–124

Fridriksson J, Kjartansson O, Morgan PS, Hjaltason H, Magnusdottir S, Bonilha L, Rorden C (2010) Impaired speech repetition and left parietal lobe damage. J Neurosci 30(33):11057–11061

Friederici AD (2009) Pathways to language: fiber tracts in the human brain. Tr Cogn Sci 13(4):175–181

Gagnon DA, Schwartz MF, Martin N, Dell GS, Saffran EM (1997) The origins of formal paraphasias in aphasics' picture naming. Brain Lang 59(3):450–472

Gao SR (1993) Aphasia. Beijing Medicine University and China Xiehe Medicine University Joint Press, Beijing

Geschwind N (1965) Disconnexion syndromes in animals and man. Brain 88(3):585–644

Goebell E, Paustenbach S, Vaeterlein O, Ding XQ, Heese O, Fiehler J et al (2006) Low-grade and anaplastic gliomas: differences in architecture evaluated with diffusion-tensor MR imaging. Radiology 239(1):217–222

Han Z, Ma Y, Gong G, He Y, Caramazza A, Bi Y (2013) White matter structural connectivity underlying semantic processing: evidence from brain damaged patients. Brain 136(Pt 10):2952–2965

Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92(1–2):67–99

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8(5):393–402

Hillis AE, Caramazza A (1995) Converging evidence for the interaction of semantic and sublexical phonological information in accessing lexical representations for spoken output. Cogn Neuropsychol 12(2):187–227

Howard D, Patterson K (1992) Pyramids and Palm Trees Test: A Test of Semantic Access from Words and Pictures. Thames Valley Test Company, Bury St Edmunds

Indefrey P (2011) The spatial and temporal signatures of word production components: a critical update. Front Psychol 2:255

Indefrey P, Levelt WJ (2004) The spatial and temporal signatures of word production components. Cognition 92(1):101–144

Lazar M, Alexander AL, Thottakara PJ, Badie B, Field AS (2006) White Matter Reorganization After Surgical Resection of Brain Tumors and Vascular Malformations. Am J Neuroradiol 27(6):1258–1271

Levelt WJ (1999) Models of word production. Tr Cogn Sci 3(6):223–232

Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS, Pandya DN (2005) Segmentation of Subcomponents within the Superior Longitudinal Fascicle in Humans: a Quantitative, In Vivo DT-MRI Study. Cereb Cortex 15(6):854–869

Maldonado IL, Moritz-Gasser S, Duffau H (2011) Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. Brain Struct Fun 216(3):1–12

Mandonnet E, Nouet A, Gatignol P, Capelle L, Duffau H (2007) Does the left inferior longitudinal fasciculus play a role in language? A brain stimulation study. Brain 130(3):623–629

Marchina S, Zhu LL, Norton A, Zipse L, Wan CY, Schlaug G (2011) Impairment of speech production predicted by lesion load of the left arcuate fasciculus. Stroke 42(8):2251–2256

Mendez MF, Benson D (1985) Atypical conduction aphasia: a disconnection syndrome. Arch Neurol 42(9):886–891

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9(1):97–113

Petrides M, Pandya DN (2009) Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. PLoS Biol 7(8):e1000170

Rolheiser T, Stamatakis EA, Tyler LK (2011) Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. J Neurosci 31(47):16949–16957

Rorden C, Karnath HO, Bonilha L (2007) Improving lesion-symptom mapping. J Cogn Neurosci 19(7):1081–1088

Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry MS et al (2008) Ventral and dorsal pathways for language. Proc Natl Acad Sci 105(46):18035–18040

Schmahmann JD, Pandya DN, Wang R, Dai G, D'Arceuil HE, Crespigny AJ, Wedeen VJ (2007) Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. Brain 130(3):630–653

Schwartz MF, Faseyitan O, Kim J, Coslett HB (2012) The dorsal stream contribution to phonological retrieval in object naming. Brain 13(Pt 12):3799–3814

Selnes OA, van Zijl PCM, Barker PB, Hillis AE, Mori S (2002) MR diffusion tensor imaging documented arcuate fasciculus lesion in a patient with normal repetition performance. Aphasiology 16(9):897–902

Shuren JE, Schefft BK, Yeh HS, Privitera MD, Cahill WT, Houston W (1995) Repetition and the arcuate fasciculus. J Neurol 242(9):596–598

Tremblay S, Shiller DM, Ostry DJ (2003) Somatosensory basis of speech production. Nature 423(6942):866–869

Vandermosten M, Boets B, Poelmans H, Sunaert S, Wouters J, Ghesquière P (2012) A tractography study in dyslexia: Neuroanatomic correlates of orthographic, phonological and speech processing. Brain 135(Pt 3):935–948

Weekes BS, Yin W, Su IF, Chen MJ (2006) The cognitive neuropsychology of reading and writing in Chinese. Lang linguist 7(3):595–617

Wernicke C (1970) The aphasic symptom complex: a psychological study on a neurological basis. Arch Neurol 22(3):280–282

Wilson SM, Galantucci S, Tartaglia MC, Rising K, Patterson DK, Henry ML et al (2011) Syntactic processing depends on dorsal language tracts. Neuron 72(2):397–403