

# White Matter Anisotropy in the Ventral Language Pathway Predicts Sound-to-Word Learning Success

Francis C. K. Wong,<sup>1</sup> Bharath Chandrasekaran,<sup>1</sup> Kyla Garibaldi,<sup>1</sup> and Patrick C. M. Wong<sup>1,2</sup>

<sup>1</sup>The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders and <sup>2</sup>Hugh Knowles Center for Clinical and Basic Science in Hearing and Its Disorders, Northwestern University, Evanston, Illinois 60208

According to the dual stream model of auditory language processing, the dorsal stream is responsible for mapping sound to articulation and the ventral stream plays the role of mapping sound to meaning. Most researchers agree that the arcuate fasciculus (AF) is the neuroanatomical correlate of the dorsal stream; however, less is known about what constitutes the ventral one. Nevertheless, two hypotheses exist: one suggests that the segment of the AF that terminates in middle temporal gyrus corresponds to the ventral stream, and the other suggests that it is the extreme capsule that underlies this sound-to-meaning pathway. The goal of this study was to evaluate these two competing hypotheses. We trained participants with a sound-to-word learning paradigm in which they learned to use a foreign phonetic contrast for signaling word meaning. Using diffusion tensor imaging, a brain-imaging tool to investigate white matter connectivity in humans, we found that fractional anisotropy in the left parietal–temporal region positively correlated with the performance in sound-to-word learning. In addition, fiber tracking revealed a ventral pathway, composed of the extreme capsule and the inferior longitudinal fasciculus, that mediated auditory comprehension. Our findings provide converging evidence supporting the importance of the ventral stream, an extreme capsule system, in the frontal–temporal language network. Implications for current models of speech processing are also discussed.

## Introduction

Brain-imaging studies have shown that both the frontal and parietal-temporal areas are engaged in speech processing (Vigneau et al., 2006). Among different neurocognitive models (Price, 2000; Friederici, 2002; Hagoort, 2005; Hickok and Poeppel, 2007), the dual stream model (Hickok and Poeppel, 2007) made explicit predictions about the anatomical bases of the temporal–frontal pathways in the language network, a dorsal stream that maps sound to articulatory representation and a ventral stream that maps sound to meaning. Although it is one of the more dominant models, recent research has called into question the role of the ventral pathway (Saur et al., 2008), which we examined in the current study.

Using diffusion tensor imaging (DTI), Glasser and Rilling (2008) reconstructed the arcuate fasciculus (AF) and demonstrated that it contains two segments, the middle temporal gyrus (MTG) segment and the superior temporal gyrus (STG) segment, that both connect to the frontal lobe. Based on a meta-analysis, they reported that the MTG termination overlapped with lexical–

semantic activations, and the STG termination overlapped with phonological activations, suggesting that the MTG segment of the AF corresponded to the ventral pathway in Hickok and Poeppel's (2007) model. Saur et al. (2008), on the other hand, provided an alternative localization of the ventral pathway. They defined nodes of a comprehension network as fMRI contrasts of activations to speech versus pseudospeech. They reported a comprehension pathway as fibers connecting the middle temporal lobe and the ventrolateral prefrontal cortex via the extreme capsule (EmC).

Saur et al. (2008) provided important insights toward the understanding of the ventral pathway by highlighting its role in comprehension, but they did not provide direct evidence that the pathway indeed interfaces semantics with phonology. The subtraction of activations to pseudospeech from activations to speech was intended to reveal activations related to comprehension and comprehension only because basic sound structure was present in both conditions, whereas meaning was distorted in pseudospeech (Wise et al., 1991; Binder et al., 2000; Röder et al., 2002).

In this study, we provided a critical examination of the interplay between phonology and semantics within the framework as laid out in Hickok and Poeppel (2007). We did so by training participants to use a foreign phonetic contrast for signaling word meaning. We trained native English speakers to use pitch patterns, changes in fundamental frequency, to contrast meanings. For example, participants learned to associate “pesh” presented with a falling pitch with a picture of a table and “pesh” presented with a rising pitch with a picture of a pencil. Using DTI, we first identified brain regions where white matter fractional anisotropy

Received Feb. 22, 2011; revised April 12, 2011; accepted May 4, 2011.

Author contributions: F.C.K.W., B.C., and P.C.M.W. designed research; B.C. and K.G. performed research; F.C.K.W., B.C., and K.G. analyzed data; F.C.K.W. and P.C.M.W. wrote the paper.

This work was supported by the National Science Foundation (BCS-0719666) and the National Institutes of Health (R01DC008333 and R21DC009652).

Corresponding should be addressed to Dr. Patrick C. M. Wong, Communication Neural Systems Research Group, Frances Searle Building 3-319, 2240 Campus Drive, Northwestern University, Evanston, IL 60208. E-mail: pwong@northwestern.edu.

B. Chandrasekaran's present address: Department of Communication Sciences and Disorders, The University of Texas at Austin, Austin, TX 78712.

DOI:10.1523/JNEUROSCI.0999-11.2011

Copyright © 2011 the authors 0270-6474/11/318780-06\$15.00/0

(FA) predicts learning success. Following Saur et al. (2008), we performed probabilistic fiber tracking to localize the pathways that mediated learning success. Our results would help resolve the competing hypotheses regarding the anatomical correlates of the ventral pathway: the MTG segment of the AF as proposed by Glasser and Rilling (2008) or the EmC as proposed by Saur et al. (2008).

## Materials and Methods

### Participants

Twenty right-handed (Oldfield, 1971) native speakers of American English (8 males and 12 females; mean age = 25.9 years, SD = 4.79) participated in the study. All participants passed a pure-tone audiometric screening at 25 dB hearing level across octaves from 500 to 4000 Hz, bilaterally, and provided their written consent before participation. The procedures were approved by the Institutional Review Board at Northwestern University. None of the participants had prior exposure to tonal languages. They were evaluated with a nonverbal IQ test (Test of Nonverbal Intelligence) (Brown et al., 1997). Their mean standard score was 119.35 (SD = 11.17). They were also evaluated with two subtests, Sound Blending (SB), a measurement of phonological awareness, and Auditory Working Memory (AWM), from the Woodcock–Johnson Test of Cognitive Abilities (Woodcock et al., 2001). Their mean percentile rank scores were 80.10 (SD = 16.71) and 85.80 (SD = 10.43), respectively.

### Sound-to-word learning program

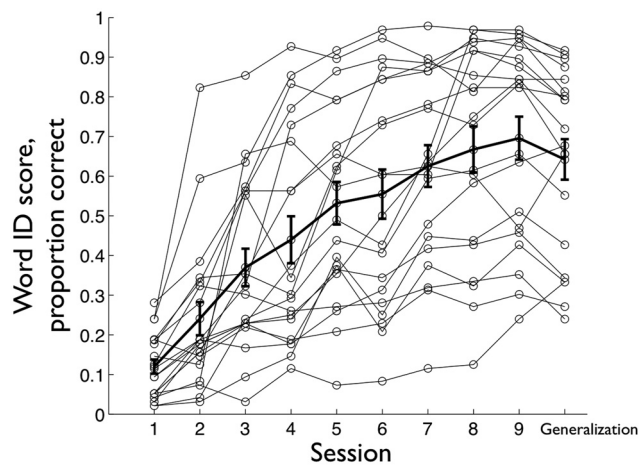
The sound-to-word learning program used in this study has been described in detail by Chandrasekaran et al. (2010) and was similar to the ones used by Wong and Perrachione (2007) and Wong et al. (2007). Participants underwent a 9 d training in which they learned to associate speech stimuli with pictures of objects presented on a computer screen. To successfully learn the sound–picture pairings, participants had to be sensitive not just to the segmental features that came naturally to them based on their native language, but also to the novel suprasegmental features, namely, changes in pitch pattern within syllables, that were not used in their native language.

**Stimuli.** As detailed in our previous study (Chandrasekaran et al., 2010), the stimuli consisted of words that were constructed based on six English monosyllabic pseudowords (“pesh,” “dree,” “nuck,” “vece,” “fute,” and “ner”), each superimposed, using the Pitch Synchronous Overlap and Add (PSOLA) method implemented in the software Praat (<http://www.praat.org>), with four pitch patterns that resembled the Mandarin Chinese level, rising, dipping, and falling tones (Fig. 1 in Chandrasekaran et al., 2010). Hence, 24 words were constructed, and each was paired with a picture designating the meaning of the word.

In total, 192 tokens were created (24 words, 8 talkers). Stimuli from four talkers (2 men and 2 women) were used as training set materials, whereas the stimuli from the remaining talkers were used in the generalization set. We used such a multitalker approach to discourage rote memorization of the acoustic inputs and promote word learning (Lively et al., 1993), the latter of which was explicitly evaluated with the Generalization Test (explained below).

**Training and testing procedures.** Each session of training lasted ~30 min; there were no more than a 2 d gap between sessions and no more than one training session per day. Each session was divided into a training phase and a Word Identification Test. In the training phase, words of the same base syllable were presented in the same block so that words in a block were minimally contrasted by pitch. In each block, each sound–picture pairing was presented with an intertrial interval of 3 s. At the end of each block, participants were tested on the words they had just learned: each sound was played and then participants had to select the correct picture from four choices. If the participant selected a wrong picture, the correct picture would be displayed briefly before the presentation of the next test item.

In the Word Identification Test, participants were presented with all the speech tokens one at a time in a pseudorandom order; they were asked to identify, untimed, each token by selecting a picture from 24



**Figure 1.** The learning curves, thin lines with circle marks, of all of the participants. The thick line with error bars, denoting the SE, shows the mean word ID scores achieved by the end of each session.

possible choices. This procedure was repeated once for each of the four talkers. No feedback was given during the test. After the Word Identification Test in the ninth session, the Generalization Test was conducted with the same procedure as in the Word Identification Test but with test items from the generalization set, that is, tokens from different talkers, instead.

### MRI data

**Image acquisition.** Diffusion-weighted images were acquired on a 3 tesla Siemens Verio with an eight-channel head coil. The diffusion-weighted volumes were achieved with 64 diffusion encoding directions with an isotropic voxel size of 2 mm<sup>3</sup> (b value = 1000 s/mm<sup>2</sup>; 55 slices; slice thickness = 2 mm; TR = 9.8 s; TE = 96 ms; FOV = 256 mm × 256 mm) plus one reference volume without diffusion weighting (b value = 0 s/mm<sup>2</sup>), which was acquired at the beginning of the sequence. The acquisition took ~11 min for each participant. T1-weighted images were acquired using an MPRAGE sequence with an isotropic voxel size of 1 mm<sup>3</sup> (160 slices; slice thickness = 1 mm; TR = 2.3 s; TE = 3.39 ms; flip angle = 9°; FOV = 256 mm × 256 mm).

**Fractional anisotropy.** The preprocessing of the diffusion data was performed using the diffusion toolbox (FDT) in FSL (Smith et al., 2004). The DTI data were first corrected for eddy currents and head motion, followed by removal of nonbrain tissues. After that, a diffusion tensor model was fitted at each voxel to compute, among other measures, FA. The FA maps created were then processed using the track-based spatial statistics routine (Smith et al., 2006), in which each individual FA map was aligned to the standard 1 × 1 × 1 mm<sup>3</sup> MNI152 space via the FMRIB58\_FA template using the nonlinear registration tool FNIRT in FSL. All coordinates reported hereafter are in MNI space. These aligned FA maps were averaged to create a mean FA map, and a thinning algorithm was applied to create a mean FA skeleton that represents the centers of all fiber bundles common to all participants. After that, each participant's aligned FA map was projected onto the skeleton such that an alignment-invariant track representation of FA values was achieved for each participant.

Note that subsequent voxelwise statistical analyses of the FA data were restricted to voxels on the skeleton, which contained 122,278 voxels (mm<sup>3</sup>) per brain.

**Statistical analysis on FA data.** To examine the unique contribution of white matter anisotropy to sound-to-word learning success, we conducted multiple regression with word identification score on the Generalization Test as the independent variable using FA and IQ as the two predictors. This regression was done using the program 3dttst++ in AFNI (<http://afni.nimh.nih.gov>). Clusters of potential significance were identified with a statistical threshold of  $t > 3.950$  (uncorrected  $p < 0.001$ ) and a cluster size larger than 20 mm<sup>3</sup>, which were then subjected to

correction for multiple comparisons based on a Monte Carlo simulation implemented in AFNI (3dClustSim).

**Probabilistic tractography.** To provide further information with regard to which language pathway mediates learning performance, we used the cluster of white matter showing significant correlation with learning success as the seed voxels in probabilistic tractography. The algorithm implemented in FSL (Behrens et al., 2003) was used in which diffusion parameters for each voxel were first estimated using a multifiber model (Behrens et al., 2007). After that, tracking was done by drawing 20,000 random samples from each seed voxel. These streamline samples started at the seed voxels and propagated through the local probability density functions of the estimated diffusion parameters.

The final product of the tractography was a connectivity map showing, for each voxel, the number of streamline samples arrived from the seed voxels. The connectivity maps of each participant were then normalized by rescaling them to a range of zero to one and thresholded to include only voxels with a normalized connectivity larger than 0.0036, the 95th percentile of the observed distribution.

The use of tractography would also provide complementary information regarding white matter connectivity. Correlation analysis using FA as a predictor is a more direct approach to look at the relationship between white matter structure and behavior. However, at voxels where fibers cross, the interpretation of FA is often hindered, whereas probabilistic fiber tracking is less prone to this problem because current tracking algorithms incorporate a multifiber model (Behrens et al., 2007).

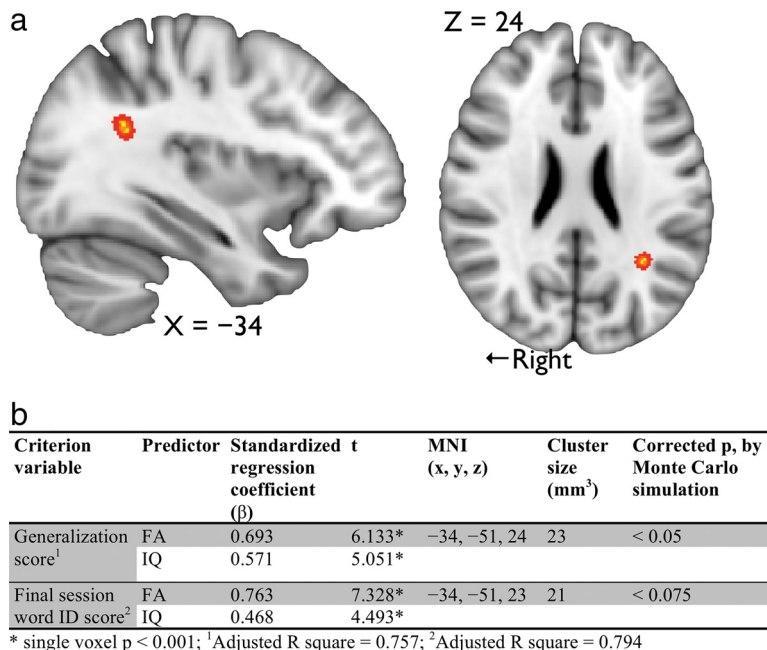
## Results

### Sound-to-word learning performance

The learning curves of the 20 participants are shown in Figure 1, where the proportion of correct word identification (word ID score) achieved in the Word Identification Test and in the Generalization Test are plotted. By the end of the training, a mean word ID score of 0.696 (SD = 0.249) on training set items was achieved and a word ID score of 0.642 (SD = 0.235) on generalization set items was achieved, hereafter referred to as the generalization score. Both the final session word ID score and generalization score correlated positively with IQ, Pearson's  $r = 0.484$  ( $p < 0.05$ ) and  $0.550$  ( $p < 0.05$ ), respectively. AWM and SB did not correlate with final session word ID score ( $r = 0.051$ ,  $p = 0.83$ , and  $r = 0.28$ ,  $p = 0.232$ , respectively) or with generalization score ( $r = 0.196$ ,  $p = 0.408$ , and  $r = 0.267$ ,  $p = 0.254$ , respectively).

### White matter anisotropy predicts learning success

Because IQ was a reliable predictor of learning success, it was included in the voxelwise multiple regression analysis to estimate the unique contribution of white matter anisotropy to learning success. Specifically, both FA and IQ were used as the predictors in the regression model, with generalization score as the criterion variable. Only one cluster, in the left parietal–temporal region (Fig. 2*a*), survived thresholding (corrected  $p < 0.05$ ) where FA positively correlated with generalization score. The same analysis was repeated with the final session word ID score as the criterion variable; although not reaching statistical significance after cor-



**Figure 2.** *a*, White matter anisotropy in the left parietal–temporal region showed significant positive correlation with generalization score ( $X = -34$ ,  $Y = -51$ ,  $Z = 24$ ; corrected  $p < 0.05$ ). The cluster was “thickened” to aid visualization. *b*, A summary of the multiple regression analysis; the first two rows: FA and IQ were predictors of generalization score; the last two rows: FA and IQ were predictors of final session word ID score.

rection for multiple comparisons (corrected  $p < 0.075$ ), a similar trend was observed.

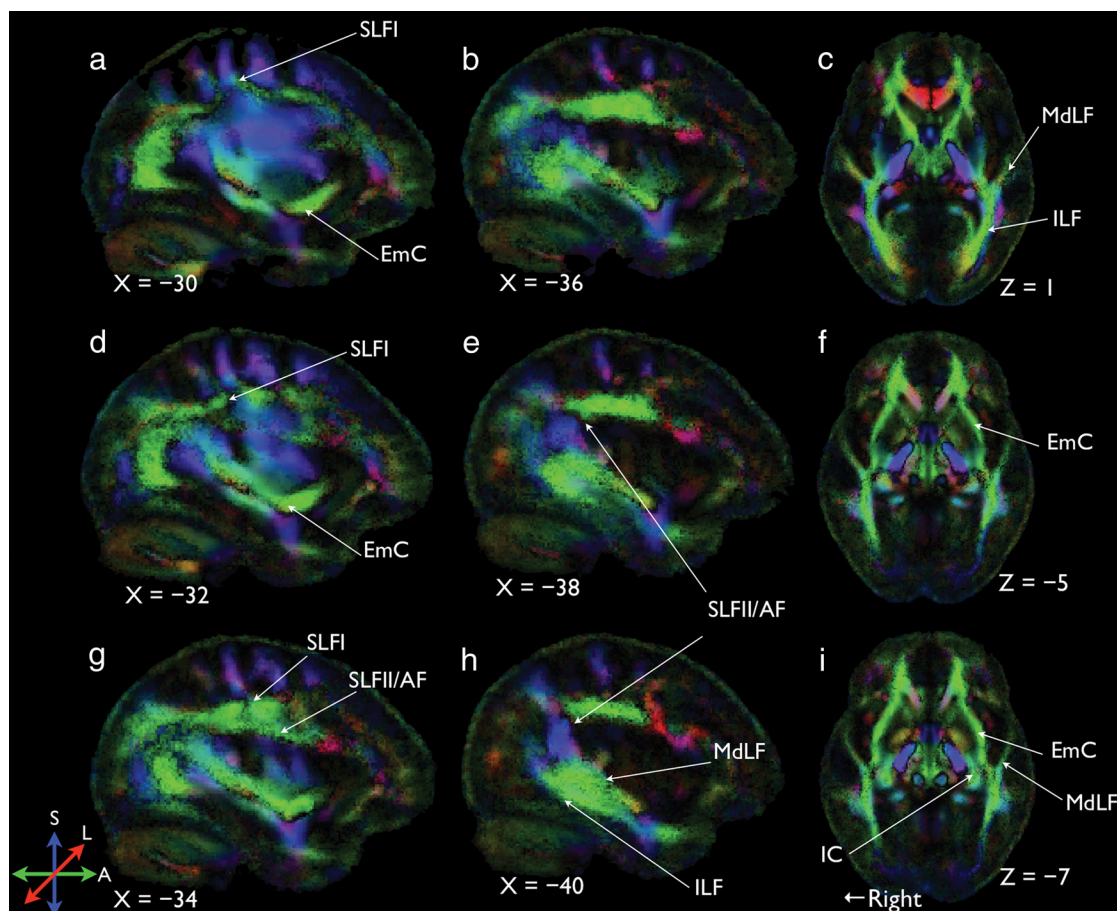
As a *post hoc* analysis, the FA values averaged across the cluster for each participant were computed and the cognitive scores were included in the regression model as predictors also. Consistent with the behavioral data, both FA and IQ predicted generalization score ( $\beta = 0.84$ ,  $p < 0.001$ , and  $\beta = 0.74$ ,  $p < 0.001$ , respectively), but neither AWM nor SB ( $\beta = -0.006$ ,  $p = 0.963$ , and  $\beta = -0.169$ ,  $p = 0.263$ , respectively) was a reliable predictor of participants' performance.

### Identification of pathways and tracking results

The mean FA maps of the 20 participants are given in Figure 3, where several long-distance tracks can be readily identified. These tracks included the superior longitudinal fasciculus I (SLFI) (Fig. 3*a,d*), the superior longitudinal fasciculus II/arcuate fasciculus (SLFII/AF) (Makris et al., 2005) (Fig. 3*e,h*), the middle inferior longitudinal fasciculus (MdLF) (Fig. 3*c,h*), the inferior longitudinal fasciculus (ILF) (Fig. 3*e,h*), and the EmC (Fig. 3*a,f*).

The results of the probabilistic tractography are summarized in Figure 4, where the mean normalized connectivity, thresholded at 0.0036, is plotted. From an operational standpoint, the pathway identified started from the seed cluster and branched out dorsally via SLFI, reaching the postcentral gyrus (Fig. 4*b*). Ventrally, it was mainly composed of the EmC, but it also made branches to the STG via the MdLF and to the MTG via the ILF (Fig. 4*c,d*). The EmC component continued in the anterior–inferior direction, running medial to the insular cortex (IC) (Fig. 4*e*), and it reached the anterior end of the insular cortex.

Note that even though the SLFII/AF was clearly visible in the FA maps (Fig. 3), it did not contribute to the frontal–temporal pathway as revealed in the tracking data. The internal capsule (Fig. 4*f*) that contains ascending fibers from the thalamus (Catani and Thiebaut de Schotten, 2008), however, was present, suggesting a putative starting point of the neural signal of the EmC system.



**Figure 3.** Mean color-coded FA maps. The principal diffusion direction was coded in RGB, where red, green, and blue code for the left–right, anterior–posterior, and superior–inferior directions, respectively. The FA was coded by the intensity of the color. Note: According to Makris et al. (2005), the AF arches around the caudal end of the Sylvian fissure (the blue segment in *h*) and extends to the lateral prefrontal cortex along with SLFII. For simplicity, we labeled such a pathway as SLFII/AF. SLFII/AF is distinguishable both by location and by endpoints from SLFI. SLFII/AF is lateral and inferior to SLFI. SLFI extends to dorsal premotor regions, but SLFII/AF does not. *a, b, d, e, g, and h*, Sagittal views at  $X = -30, -36, -32, -38, -34$ , and  $-40$ , respectively; *c, f, and i*, axial views at  $Z = 1, -5$ , and  $-7$ , respectively.

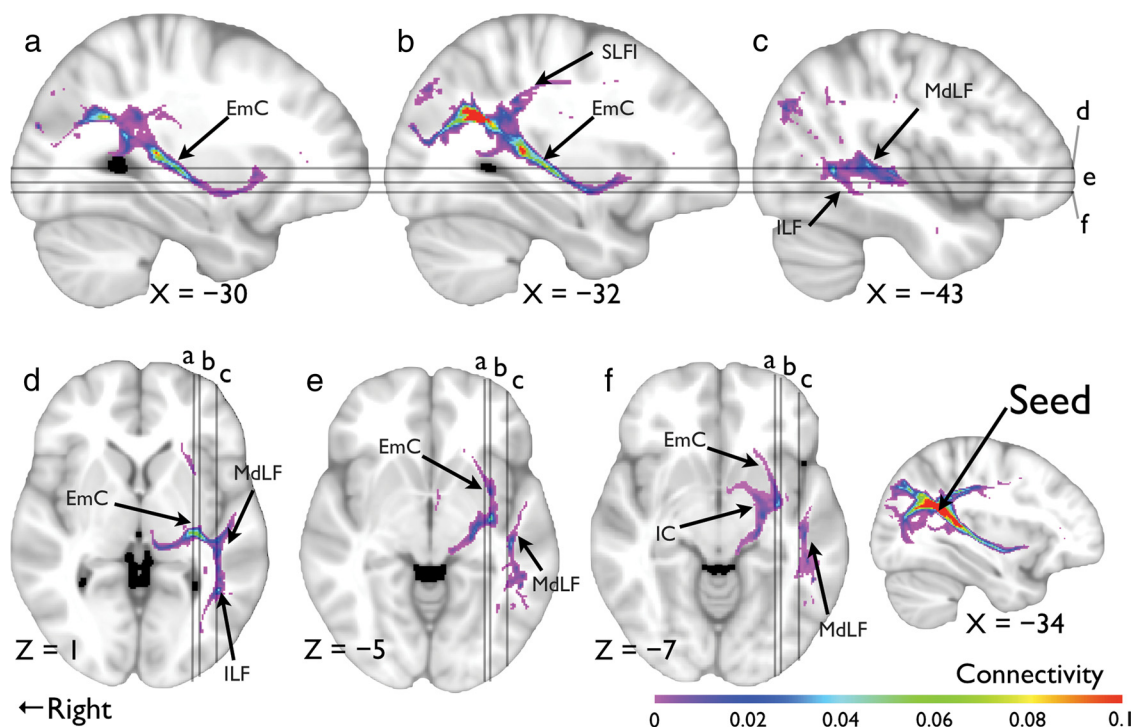
## Discussion

In this study, we report that white matter anisotropy in the left parietal–temporal region predicts sound-to-word learning success. Anisotropy of water diffusion in white matter has been argued to be due to the dense packing of intact axons that restrict water diffusion perpendicular to the axons (Beaulieu, 2002). The positive correlation between FA and generalization score therefore suggests that the denser the packing of the white matter in that region, the more successful one will be. Voxels in the left parietal–temporal were used as the seed for probabilistic tractography to infer the white matter pathways that mediate sound-to-word learning success, and hence the putative pathways interfacing acoustic processing with semantic processing. The tractography results reveal an extreme capsule system that subserves the function of mapping sound to meaning. This ventral sound-to-meaning pathway is composed of an MdLF branch running along the STG, an ILF branch running along the MTG, a short segment of the SLFI and the EmC, and a long-distance fiber connecting the parietal–temporal region with the inferior frontal region. Thus, our results converge with those of Saur et al. (2008, 2010) and collectively provide evidence for the importance the ventral pathway (Weiller et al., 2009), but not the dorsal pathway (Friederici, 2009; Brauer et al., 2011), in comprehension. Nevertheless, the importance of the ventral stream, an extreme capsule system, in the frontal–temporal language network had been pro-

posed by Wernicke more than 130 years ago (see Catani and Mesulam, 2008, for review).

It is worth noting that although we followed Saur et al. (2008, 2010) in our implementation of the DTI methodology in investigating the functional role of white matter pathways, our data complement their findings in a critical way. Specifically, we argue that they did not demonstrate the interaction between phonological and semantic processing. The nodes that they defined for tractography analysis reflected either phonological or semantic processing in isolation. In Saur et al. (2008), for instance, nodes for a comprehension network were defined as the subtraction of activations to pseudospeech from activations to speech with the assumption that such a contrast would reveal activations related to comprehension but not activations related to phonological processing. A similar approach was used in Saur et al. (2010) to reveal a phonological network using seeds defined as fMRI contrast between pseudospeech and reversed speech. One important function of a pathway in a brain network is to provide an interface between two or more processes to function hand-in-hand for a unified goal, that is, to comprehend speech. In this study, we attempted to directly investigate the relationship between phonological and semantic processing by using a sound-to-word learning paradigm.

In light of Hickok and Poeppel's (2007) model as well as findings from our previous studies that have used similar sound-to-



**Figure 4.** Probabilistic tractography results showing connections to EmC, MdLF, IC, and ILF from the left parietal-temporal region (the seed). *a, b, c*, Sagittal views at  $X = -30$ ,  $-32$ , and  $-43$ , respectively; *d, e, f*, axial views at  $Z = 1$ ,  $-5$ , and  $-7$ , respectively.

word learning paradigms, we interpret our findings as follows. We previously reported that sound-to-word learning performance was positively correlated with pitch perception ability (Wong and Perrachione, 2007) as well as with the size of the left Heschl's gyrus (Wong et al., 2008). White matter density in the left Heschl's gyrus was also found to be higher in fast learners, compared with slow learners, in an auditory nonword learning paradigm (Golestani et al., 2007). Together, these findings highlighted the importance of primary acoustic analysis in learning a new phonemic contrast.

With the Heschl's gyrus as the first station of the cortical auditory language processing network, it has been suggested that neural signals propagate along the superior temporal sulcus (STS) for higher level processing, as evidenced by an increase in acoustic invariance in the anterior (aSTS) and posterior STS (pSTS) (Okada et al., 2010; Peelle et al., 2010). The pSTS, in particular, was proposed to be the phonological network according to Hickok and Poeppel (2007). The MdLF branch in our tracking data (Fig. 4c) seems to play a role in propagating signal from the Heschl's gyrus downstream to aSTS and pSTS.

The white matter in the parietal-temporal region (Fig. 2a) appears to be a hub in the EmC system because connections from the MdLF make branches there via the SLFI to postcentral gyrus, via the ILF to the MTG and via the EmC to the inferior frontal region. Both the ILF branch and the EmC branch contribute to the ventral sound-to-meaning pathway. The ILF branch is the anatomical basis of Hickok and Poeppel's (2007) prediction about the ventral stream. They suggested that it is a pathway running ventral to the Sylvian fissure connecting the phonological network in the middle to posterior portion of STS with the inferior temporal region such as the posterior MTG. The ILF branch revealed in our tractography data is also in partial agreement with the findings of Glasser and Rilling (2008), in which they argued that the MTG segment of the AF constitutes the

ventral stream subserving comprehension. However, our data suggest that this ventral branch is subserved by the ILF rather than by the AF.

The EmC branch, however, is not included explicitly in Hickok and Poeppel's (2007) model. Although there is no information from our data to locate the frontal terminations of the EmC, Frey et al. (2008) suggested that this long-association fiber system connects the STG with BA 45, an area that has been implicated in semantic processing (Friederici, 2002; Hagoort, 2005). The semantic function of the EmC is in agreement with work by Saur et al. (2008, 2010), in which a comprehension network was found to be composed of the EmC, MdLF, and ILF.

Our data are also consistent with the dual auditory processing model proposed by Rauschecker and Scott (2009). They suggested that speech production and perception share a common computational framework, operate together, and support each other. In their model, this production-perception interdependence is achieved by a "forward mapping" projection and an "inverse mapping" projection. They suggested that the forward mapping projection from motor preparatory networks (inferior frontal and premotor cortex) interfaces with the inverse mapping projection from auditory cortex via the inferior parietal lobule (IPL). It is interesting that this is the same region that we refer to as a "hub" in the EmC system, as per our FA and tractography results. The SLFI segment (Fig. 4b) that was identified seems to be the putative anatomical correlate of the proposed connection between premotor cortex and the IPL in the model of Rauschecker and Scott (2009). It is important to note that their model was based on neuroanatomical data from nonhuman primates, and our study potentially provides support for the Rauschecker and Scott model with human data.

Our previous studies suggested that primary acoustic processing is important to successful learning of a new phonemic contrast. The correlation analysis conducted here was driven by word

identification performance. Although representing the state of the art in DTI analysis, we do acknowledge the limitation of the current study that although DTI allows us to investigate the connectivity in a living brain, it can provide only anatomical information (Duffau, 2008). The functional interpretation of white matter connectivity still relies on indirect evidence (Friederici, 2009). Nevertheless, our findings, together with those of Saur et al. (2008, 2010), provide critical comprehensive evidence that there exists a dual ventral pathway for speech comprehension, one subserved by the inferior longitudinal fasciculus and the other subserved by the extreme capsule. The differential roles played by these two streams, however, remain a subject for further research.

## References

- Beaulieu C (2002) The basis of anisotropic water diffusion in the nervous system—a technical review. *NMR Biomed* 15:435–455.
- Behrens TE, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, Matthews PM, Brady JM, Smith SM (2003) Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 50:1077–1088.
- Behrens TE, Berg HJ, Jbabdi S, Rushworth MF, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34:144–155.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and non-speech sounds. *Cereb Cortex* 10:512–528.
- Brauer J, Anwander A, Friederici AD (2011) Neuroanatomical prerequisites for language functions in the maturing brain. *Cereb Cortex* 21:459–466.
- Brown L, Sherbenou RJ, Johnsen SK (1997) TONI-3: Test of nonverbal intelligence, Ed 3. Austin, TX: Pro-Ed.
- Catani M, Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44: 953–961.
- Catani M, Thiebaut de Schotten M (2008) A diffusion tensor imaging tractography atlas for virtual *in vivo* dissections. *Cortex* 44:1105–1132.
- Chandrasekaran B, Sampath PD, Wong PCM (2010) Individual variability in cue-weighting and lexical tone learning. *J Acoust Soc Am* 128:456–465.
- Duffau H (2008) The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. *Neuropsychologia* 46:927–934.
- Frey S, Campbell JS, Pike GB, Petrides M (2008) Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *J Neurosci* 28:11435–11444.
- Friederici AD (2002) Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6:78–84.
- Friederici AD (2009) Allocating functions to fiber tracts: facing its indirectness. *Trends Cogn Sci* 13:370–371.
- Glasser MF, Rilling JK (2008) DTI tractography of the human brain's language pathways. *Cereb Cortex* 18:2471–2482.
- 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.
- Hagoort P (2005) On Broca, brain, and binding: a new framework. *Trends Cogn Sci* 9:416–423.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
- Lively SE, Logan JS, Pisoni DB (1993) Training Japanese listeners to identify English /r/ and /l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. *J Acoust Soc Am* 94: 1242–1255.
- Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS Jr, Pandya DN (2005) Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, *in vivo*, DT-MRI study. *Cereb Cortex* 15:854–869.
- Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences JT, Hickok G (2010) Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. *Cereb Cortex* 20:2486–2495.
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 9:97–113.
- Peelle JE, Johnsrude IS, Davis MH (2010) Hierarchical processing for speech in human auditory cortex and beyond. *Front Hum Neurosci* 4:51.
- Price CJ (2000) The anatomy of language: contributions from functional neuroimaging. *J Anat* 197:335–359.
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724.
- Röder B, Stock O, Neville H, Bien S, Rösler F (2002) Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: a functional magnetic resonance imaging study. *Neuroimage* 15:1003–1014.
- 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.
- Saur D, Schelter B, Schnell S, Kratochvil D, Küpper H, Kellmeyer P, Kümmerer D, Klöppel S, Glauche V, Lange R, Mader W, Feess D, Timmer J, Weiller C (2010) Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage* 49:3187–3197.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazky 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.
- 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.
- Vigneau M, Beaucois V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N (2006) Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30:1414–1432.
- Weiller C, Musso M, Rijntjes M, Saur D (2009) Please don't underestimate the ventral pathway in language. *Trends Cogn Sci* 13:369–370; 370–371.
- Wise R, Chollet F, Hadar U, Friston K, Hoffner E, Frackowiak R (1991) Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain* 114:1803–1817.
- Wong PCM, Perrachione TK (2007) Learning pitch patterns in lexical identification by native English-speaking adults. *Appl Psycholinguist* 28:565–585.
- Wong PCM, Perrachione TK, Parrish TB (2007) Neural characteristics of successful and less successful speech and word learning in adults. *Hum Brain Mapp* 28:995–1006.
- Wong PCM, Warrier CM, Penhune VB, Roy AK, Sadehh A, Parrish TB, Zatorre RJ (2008) Volume of left Heschl's gyrus and linguistic pitch learning. *Cereb Cortex* 18:828–836.
- Woodcock RW, McGrew KS, Mather N (2001) Woodcock–Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside Publishing.