Fusiform Gyrus Laterality in Writing Systems with Different Mapping Principles: An Artificial Orthography Training Study

Elizabeth A. Hirshorn¹, Alaina Wrencher¹, Corrine Durisko¹, Michelle W. Moore¹,² and Julie A. Fiez¹

Abstract

- Writing systems vary in many ways, making it difficult to account for cross-linguistic neural differences. For example, orthographic processing of Chinese characters activates the mid-fusiform gyri (mFG) bilaterally, whereas the processing of English words predominantly activates the left mFG. Because Chinese and English vary in visual processing (holistic vs. analytical) and linguistic mapping principle (morphosyllabic vs. alphabetic), either factor could account for mFG laterality differences. We used artificial orthographies representing English to investigate the effect of mapping principle on mFG lateralization. The fMRI data were compared for two groups that acquired foundational proficiency: one for an alphabetic and one for an alphasyllabic artificial orthography. Greater bilateral mFG activation was observed in the alphasyllabic versus alphabetic group. The degree of bilaterality correlated with reading fluency for the learned orthography in the alphasyllabic but not alphabetic group. The results suggest that writing systems with a syllable-based mapping principle recruit bilateral mFG to support orthographic processing. Implications for individuals with left mFG dysfunction will be discussed.

INTRODUCTION

The role of a proposed visual word form area (VWFA) has been a widely investigated topic over more than a decade. The standard profile of this functionally defined area in the mid-fusiform gyrus (mFG) is that it shows greater selectivity for words or alphabetic strings than other visual stimuli (McCandliss, Cohen, & Dehaene, 2003; Cohen et al., 2002), but the underlying mechanisms that lead to this selectivity remain under debate. For instance, it is unclear whether the specialization of this region is driven by visuoperceptual (Dehaene & Cohen, 2011) or linguistic/phonological processing demands (Price & Devlin, 2011). Nevertheless, its importance in the reading network has been made clear from a diverse set of evidence. For example, the typical selectivity of this region for words is seen only after reading instruction, dyslexic readers show atypical engagement of the mFG (Richlan, Kronbichler, & Wimmer, 2009; Vigneau et al., 2006; Price & Devlin, 2003; Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002), and acquired damage to the mFG is associated with acquired alexia (Leff, Spitsyna, Plant, & Wise, 2006; Warrington & Shallice, 1980) or “word blindness.” Thus, it remains important to understand more precisely the factors that influence the specialization of this region, because this should enhance our understanding of normal and disordered reading development, and it may lead to improved strategies for remediation.

An infrequently studied question is whether the role of the mFG differs across writing systems, and if so why. During reading tasks, the mFG tends to be left-lateralized for alphabetic languages (Nelson, Liu, Fiez, & Perfetti, 2009; Price & Devlin, 2003; Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002), but the majority of findings obtained through studies of English. In contrast, bilateral activation of the mFG has been observed for Chinese (Nelson et al., 2009; Liu, Dunlap, Fiez, & Perfetti, 2007; Tan et al., 2001). Some past researchers have attributed the differences in laterality across the two writing systems to the fact that Chinese involves greater visual spatial processing demands (Liu et al., 2007; Tan et al., 2001). To test the effect of visual demands on fusiform engagement, a recent study leveraged the fact that faces tend to be processed in the right hemisphere. This motivated the creation of an artificial orthography in which faces were used as letters to represent the consonant and vowel sounds of English. This face-alphabet (termed “FaceFont”) was compared with “KoreanFont,” a linguistically equivalent alphabetic system in which Korean letters were used to represent English phonemes (Moore, Durisko, Perfetti, & Fiez, 2014) (Figure 1). Both were transparent alphabetic systems, in which 35 letter–sound pairs were used to represent all of the consonant and vowel sounds of English. The authors reasoned that if
fusiform laterality is sensitive to visuoperceptual processing demands, then FaceFont should engage the right hemisphere to a greater extent than KoreanFont. They found that a region in the left mFG, within the typical territory of the VWFA, responded to both FaceFont and KoreanFont words more strongly in trained versus untrained participants, whereas no differences were observed in the right mFG. These results suggest that left-lateralized mFG processing of orthographic stimuli is not restricted to stimuli with particular visual-perceptual features.

Several research groups have used artificial orthographies to explore an alternative hypothesis for the laterality of the mFG: namely, that the mapping principle of a writing system plays a role in determining the laterality of orthographic processing. Instead of varying the visual characteristics of an alphabetic writing system, as was done by Moore, Durisko, et al. (2014), two groups kept the artificial orthography the same but varied the method of instruction. Some participants were explicitly taught the letter–sound correspondences of an alphabet, whereas others were taught whole-word correspondences between printed and spoken word forms (Yoncheva, Wise, & McCandliss, 2015; Mei et al., 2013; Yoncheva, Blau, Maurer, & McCandliss, 2010). Results from these studies suggest that both the left and right mFG contribute to orthographic processing when an alphabetic writing system is taught as a logographic (whole-word) system.

Moore and colleagues, in a second study that used faces as orthographic stimuli, also investigated the impact of mapping principle (Moore, Brendel, & Fiez, 2014). They did so by directly varying the mapping principle of two different artificial proto-orthographies. For an alphabetic system, 10 face graphemes were mapped onto 10 English phonemes, which could be combined to create words. For a syllabic system, 15 face graphemes were mapped onto English syllables (which were also words, such as may, be, four), which could be combined to create English words (e.g., maybe, before). They tested the ability of a patient with a left occipitotemporal lesion (due to a stroke) to acquire the two orthographies. Importantly, the patient’s lesion encompassed the typical territory of the VWFA, and she exhibited the hallmark symptoms of acquired alexia, a loss of fluent word recognition that has been associated with damage to the VWFA. The patient struggled with learning the phoneme–grapheme pairings, but she was able to learn all of the syllable–grapheme pairings and use them to decode novel words. This result provides further evidence that the grain size of the orthographic mapping principle influences the laterality of the neural underpinnings required for reading.

The current study also uses an artificial orthography approach to test the mapping principle account of mFG lateralization during reading. However, it goes beyond the prior work through the use of an artificial orthography in which a corpus of 375 syllable-grapheme mappings can be used to represent any spoken English word. Face images are used as the component graphemes in the system, which is termed “Faceabary” (Hirshorn & Fiez, 2014). The Faceabary system can be regarded as either a syllabic system, because each of the 375 face graphs is mapped onto a single English syllable or as an alphasyllabic system, because both consonant and vowel information can be identified within the graphs (Figure 1). This is because different face identities systematically map to different consonants, and different facial expressions systematically map to different vowels. These mappings are not explicitly taught to participants, but participants do become aware of this general design principle, and the alphasyllabic structure likely facilitates participants’ ability to master the syllable-grapheme correspondences in the Faceabary system. A few simple decoding rules specify how a sequence of represented syllables can be blended together to produce the phonological form of a word printed in Faceabary. For instance, a “vowel-dropping” rule is used to blend together two consonant–vowel (CV) Faceabary graphs to represent a CVC English word. Thus, the graphs for /kæ/ and /tə/ can be combined to represent the word “cat.” Using the graphs in the Faceabary system and its decoding rules, it is possible to represent any English word.

In this study, participants were trained to read using the Faceabary system. They were first trained to decode words and then to read whole texts, in an effort to mimic natural reading acquisition. The learning of the syllable-based Faceabary system was compared with the learning of the phoneme-based (alphabetic) FaceFont system reported in the study by Moore, Durisko, et al. (2014). Functional neuroimaging (fMRI) was used to provide measures of mFG lateralization following training. The mapping principle account of mFG lateralization predicts that a left-lateralized training effect should be observed in the fusiform gyri for the alphabetic system, while a bilateral pattern should be observed for the syllabic system.

**Figure 1.** Example stimuli for fMRI experiment. Each group exclusively saw face-words in their learned artificial orthography, in addition to identical KoreanFont and pattern stimuli.

**METHODS**

**Participants**

All participants were native English speakers who completed an initial screening in which they reported no history of hearing or vision issues, learning or reading...
difficulties, drug or alcohol abuse, mental illness, or neurological problem. Additionally, participants were screened for MRI contraindications (e.g., ferromagnetic material in or on body, not right-handed, claustrophobic, pregnant, etc.). All participants provided informed consent and were compensated for their time.

**FaceFont**

Twelve participants (four men) completed the 2-week FaceFont training protocol and a subsequent posttraining behavioral and neuroimaging session ($M$ age = 21.8 years, $SD = 2.0$). Training study participants were recruited through fliers posted around the University of Pittsburgh campus (see Moore, Durisko, et al., 2014).

**Faceabary**

Fifteen participants (six men) completed the 3-week Faceabary training protocol, in addition to pre- and post-training behavioral and neuroimaging sessions ($M$ age = 20.6 years, $SD = 1.7$). Participants were recruited from a database of individuals who had participated in previous behavioral studies and had indicated that they would be interested in participating in future studies.

**Writing Systems**

Each participant was assigned to learn either the FaceFont or the Faceabary writing system. Both of these artificial orthographies use faces from the NimStim set (Tottenham et al., 2009) as graphemes, but the faces are mapped to English phonology using different mapping principles.

**FaceFont**

FaceFont is a transparent alphabetic system with a one-to-one grapheme (face)-to-phoneme correspondence mapping principle. Thirty-five face-sound pairs are used to represent all sounds in English. There are five exceptions in which a single grapheme (i.e., a single face) represents two similar sounds (e.g., /ɔ/ in hawk and /a/ in hot).

**Faceabary**

The Faceabary system consists of 375 unique grapheme/faces that represent a syllable. There is internal consistency such that all face identities represent the consonant component of the syllable and the displayed facial emotion represents the vowel component. We chose the identities for Faceabary from the NimStim that included the full range of expressions. There were 13 total expressions with open and closed mouth. This resulted in 21 faces used for simple CV graphs, and 16 more for more complex graphs (see Syllable–Grapheme Mapping Training section).

**Behavioral Training Procedure**

**FaceFont Training**

FaceFont-trained participants completed 1- to 2-hr training sessions consisting of three components: grapheme (face phoneme mapping) training (Session 1), word level training (Sessions 2–5), and story level training (Sessions 6–9). Progress was monitored at the end of each word level and story level training session with a single-word reading test. Training was followed by a final session on the 10th day. During this final session, participants read stories that were transcribed from a standardized reading test designed to assess reading fluency and comprehension (Wiederholt & Bryant, 2001). They also completed an fMRI session to probe for the neural basis of FaceFont reading. No printed English was used except for the initial basic instructions of a task. For details beyond those provided below, see Moore, Durisko, et al. (2014).

**Phoneme–grapheme mapping training.** In Session 1, participants completed phoneme training. Using the E-Prime computer program for psychological experiments (Psychology Tools, Inc., Pittsburg, PA; Schneider, Eschman, & Zuccolotto, 2002), a grapheme would appear on the computer screen, and participants pressed a button to elicit the auditory presentation of the associated phoneme. Because the focus was on the participants achieving mastery of all 35 grapheme–phoneme pair associations, the participants could spend as much time as they wanted on each grapheme, and each associated phoneme could be played unlimited number of times before the participant advanced to the next grapheme. After all 35 pairs were presented in random order, the cycle was repeated four more times, for a total of five cycles of individually paced grapheme–phoneme learning.

**Phoneme test.** Participants took a phoneme test after they completed the phoneme training. Graphemes appeared on the computer screen, and participants were asked to say aloud the phoneme associated with each grapheme. Two cycles were administered with random order of grapheme presentation for 70 total items on the test. Participants were required to score 90% or better on the phoneme test. If they did not meet this criterion, the examiner reviewed their specific errors, and they repeated both the phoneme training and phoneme testing as many times as needed until criterion was reached or until session duration was exceeded. All participants passed the phoneme test in three or fewer attempts.

**Word level training.** In Sessions 2–5, participants completed word level training in which they read 400 one-syllable words, two to four phonemes in length. Words were presented in random order using E-Prime. The same 400 words were used in each session to facilitate fluency for reading through repetition. Participants were encouraged to attempt to read the word when it appeared
on the screen but had the option to hear any of the individual phonemes or to hear the whole word. For example, one of the training words was “beef,” consisting of phonemes /b/, /ə/, /f/. Using a key press, participants could play any of the three phonemes individually or could play the entire word, if necessary.

After completing each word level training session (Sessions 2–9), participants took single-word reading tests on the computer consisting of “old” words (words that were in the word training), “new” words (words that were not in the word training task), and nonwords (15 each). Participants were presented with items one at a time and asked to read each one aloud as quickly and as accurately as possible. Results are reported in Moore, Durisko, et al. (2014).

Faceabary

Faceabary-trained participants completed training similar to FaceFont-trained participants, but it extend over 3 weeks because of the much larger set of face–syllable mappings to be learned. Additionally, rather than learning all of the graph-sound mappings before commencing word level training, the large corpus of Faceabary graphs was divided into subsets, and participants progressed through multiple cycles of grapheme and interleaved word training to give the following: grapheme (face–syllable mapping) training (Sessions 1–9), word level training (Sessions 1–10), and story level training (Sessions 11–15). Another difference was that Faceabary training was preceded by an fMRI session. Faceabary training was similarly followed by an imaging session on the last (15th) day.

Syllable–grapheme mapping training. The first nine sessions were used for learning the mappings between face graphemes and syllables. The basic procedure was the same each day as in the FaceFont training: Participants would see a face grapheme on the screen, then participants pressed a button to elicit the auditory presentation of the associated phoneme. Participants could spend as much time as they wanted on each grapheme, and each associated syllable could be played unlimited number of times before the participant advanced to the next grapheme. During the first week of training, participants learned all simple CV syllables. Because of the large number of graphemes to be learned and the internal structure of Faceabary, training each day focused on two to three vowels and all of their respective consonant pairings (e.g., Day 1 focused on C + /i/ and C + /a/ syllables for a total of 42 total syllables). On Day 2, participants learned CV syllables that focused on /a/, /o/; Day 3 on /a/, /æ/, /ɛ/; Day 4 on /ou/, /æ/, and /ɛ/; Day 5 on /æ/, /a/, and /ɔ/; Day 6 on /i/, /i/, /k/, and so on, and then /p, /t, /k, /g/, and so on. After all pairs were presented, the cycle was repeated four more times, for a total of five cycles of individually paced grapheme–syllable learning. After Day 1, the session started with a recap of previously learned face graphemes before learning new ones. During the second week (Days 6–9), participants learned non-CV graphemes. On Day 6 they learned CCV graphemes, on Day 7 they learned VC graphemes, on Day 8 they learned VCC graphemes, and on Day 9 they learned faces that represented grammatical markers.

Syllable tests. Participants took a syllable test every day after they completed the grapheme–syllable mapping training. Graphemes appeared on the computer screen in random order, and participants were asked to say aloud the syllable associated with each grapheme. Participants were required to score ≥75% or better. If they did not meet this criterion, the examiner reviewed their specific errors, and they repeated both the syllable training and syllable testing as many times as needed until criterion was reached or until session duration was exceeded. All participants reached criterion within three training-testing cycles.

Word level training. Each day after syllable training, participants completed word level training in which they read 50 words that focused on newly learned graphemes but could draw upon previous learned graphemes as well. On the first day, participants were instructed on decoding rules involving “dropping” the schwa at the ends of words (e.g., /ni/ + /ta/ = neat) or between successive consonants (e.g., /pa/ + /leu/ = play). Participants were encouraged to attempt to read the word when it appeared on the screen but received feedback from the experimenter when they incorrectly pronounced a word or struggled with a certain grapheme. After completing all grapheme training, on Day 10 participants reviewed all words (n = 450) that they had read in previous days. After completing the syllable–grapheme mapping training and word level training, participants took single-word reading tests on the computer (Sessions 11–15), similar to that of FaceFont participants. Stimuli consisted of “old” words (words that were in the word training), “new” words (words that were not in the word training task), and nonwords. Participants were presented with items one at a time and asked to read each one aloud as quickly and as accurately as possible. Results are outside the scope of the current article and are not reported.

Story Level Training (FaceFont and Faceabary)

After grapheme and word training, all participants spent the last week of training (Week 2 for FaceFont, Week 3 for Faceabary) reading stories in their respective writing systems. Each day, they read 10 early reader stories from the “Now I’m Reading!” series (Gaydos, 2003) that were transcribed into the training fonts, beginning with 10 Level
and more complex stories (up to Level 4). Story level reading performance was measured by words read per minute (WPM), calculated by the total number of words in each story divided by the time to complete the passage.

During the final session for both groups (Session 10 for FaceFont, 15 for Faceabary), participants read the first six stories from Form A of the Gray Oral Reading Test-4 (GORT-4; Wiederholt & Bryant, 2001), which had been transcribed into their respective training font. Participants were asked to read each story aloud as quickly and accurately as possible. The stories were administered and scored according to the standardized test protocol. Standardized GORT scores were not used because the stories had been transcribed into our experimental fonts and participants in our study exceeded the maximum age on which normative data have been collected for this test. Administration of the multiple-choice comprehension questions deviated from the protocol in that the questions/answers were only read aloud to the participants (whereas in the standardized protocol, the items are presented both visually and aurally). As the transcribed materials deviate significantly from the published GORT, we will therefore refer to this task and its measures as the Face Oral Reading Test (FORT). Raw FORT scores for the following were obtained: accuracy, time, and comprehension. In addition to the raw scores, we computed words per minute (WPM) for each story as a normalized measurement of time/reading rate across the stories that varied in length (Table 3). Lastly, we calculated a z-score for each individual measure (accuracy, WPM [time], and comprehension), and then created a composite mean FORT z-score across the three measurements for all participants. The mean FORT z-score was calculated using the z-score for comprehension and WPM, where positive scores reflected greater skill. Because a positive z-score on the deviations from print (accuracy) measure reflected poorer performance, we used their accuracy score times −1, in order for positive scores to reflect better performance in all three measurements. The mean FORT z-score was used for further analyses.

**fMRI Experimental Procedure**

Training study participants were shown three types of stimuli—face-words (in FaceFont or Faceabary), Korean-font words used in a previous study where Korean characters served as graphemes (Moore, Durisko, et al., 2014), and patterns (used as a baseline). Participants were instructed to stay alert and passively view all stimuli. Although they were not explicitly told to read the face-words, they all reported trying to do so during the scans. Face-words on average were 2.16 graphs long in Faceabary and 3.25 long in FaceFont. Patterns were matched for length to the face-words in both groups (Figure 1). To ensure that participants were not falling asleep, we had visual monitoring and could see eye-blinks throughout the scan. They completed two functional runs designed identically. The two runs had the same block design format, but the order of the blocks and the stimuli within each block were randomized across the two runs for each participant. Each run contained 21 20-sec epochs, with seven epochs of each stimulus type presented in random order. Each epoch consisted of 10 2-sec trials using the same stimulus type for each trial. Within a trial, a stimulus item was presented for 1500 msec, followed by a “+” for 500 msec. The next trial immediately followed with another stimulus item of the same type for 1500 msec and “+” for 500 msec, and so on. There was no pause between epochs.

Within this paradigm, each training study participant passively viewed 140 English words printed in their specific training font. Both the FaceFont group and the Faceabary group used the same word list. Words were one syllable, two to four phonemes in length. No word was repeated within the fMRI session, nor did the words overlap with the items used in the word training set from the behavioral sessions. Words were presented in random order across both runs. Faceabary participants completed this imaging protocol before and after training, whereas FaceFont participants were only scanned after training.

**fMRI Data Analysis**

A series of preprocessing steps were conducted before data analysis using the integrated NeuroImaging Software package (NIS 3.6; Fissell et al., 2003) to correct for artifacts and movement and to account for individual differences in anatomy. Images were reconstructed and then corrected for subject motion with Automated Image Registration (AIR 3.08; Woods, Cherry, & Mazziotta, 1992). For runs in which head motion exceeded 4 mm or 4° in any direction, data from the beginning of the epoch in which the head movement occurred through the end of the run were not used in the analysis. The images were then corrected to adjust for scanner drift and other linear trends within runs. The structural images of each participant were stripped to remove the skull and coregistered to a common reference brain, chosen from among the
participants (Woods, Mazzotta, & Cherry, 1993). Functional images were transformed into the same reference space, normalized by a mean scaling of each image to match global mean image intensity across participants, and smoothed using a 3-D Gaussian filter (8 mm FWHM) to account for anatomical differences between participants. Finally, images were converted into Talairach space (Talairach & Tournoux, 1988). Specific cerebellar lobules were designated using a cerebellar atlas (Schmahmann, Doyon, Toga, Petrides, & Evans, 2000).

**Group comparisons in literature-derived fusiform ROIs.** Because FaceFont participants did not complete a localizer task, an ROI approach was used to investigate neural activity in the mFG at or near the prototypical site of the VWFA and a potential right hemisphere homologue. Specifically, we used the $x$-$y$-$z$ peak coordinate from Bolger, Perfetti, and Schneider (2005) ($-45-57-12$ and right homologue) that was sampled from different writing systems.

**Group comparisons in whole brain.** NIS 3.6 was used to analyze the fMRI data. We performed a whole brain, voxel-wise $2 \times 2$ ANOVA with Training group (FaceFont vs. Faceabary) and Stimulus (Face-words vs. Patterns). Because there were no pauses between epochs during task administration, the first four 2-sec time periods (T1–T4) of each epoch were not included in the analysis to allow the carryover BOLD response to the stimuli in the prior epoch to return to baseline; therefore, only T5–T10 were included in the ANOVA. AFNI 3dClustSim (Cox, 1996) was used to identify a voxel-wise significance threshold that was corrected to $p < .05$ (uncorrected at $p < .01$, cluster size of 42 voxels).

**Faceabary pre- versus postscan comparison.** The same analysis approach was used to compare the Faceabary results pre- versus posttraining, except the $2 \times 2$ ANOVA was within the Faceabary group, and the factors were Training time point (Pre vs. Post) and Stimulus type (Face-words vs. Patterns). Results are reported at $p < .05$ corrected.

**Faceabary pre- versus postscan functional connectivity.** An analysis was run in the Faceabary group comparing pre- and postscan functional connectivity to determine if different neural networks were becoming engaged after training. Seeding took place in a functionally derived right mFG ROI that showed greater activation for face-words than patterns in the Faceabary group compared with the FaceFont group during the posttraining scan. We ran analyses on the time series of only the face-word blocks pre- and posttraining using condition-specific psychophysiological interactions (O’Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). The mean time series for “face-word” trials was extracted for each participant for the ROI seed region. A simple functional correlation analysis was conducted using a standard procedure within AFNI (afni.nimh.nih.gov/sscc/gangc/SimCorrAna.html). Time series were used as a regressor in single-subject analyses using standard GLM methods, including nuisance variables that modeled head motion, white matter, CSF, and the global mean. Differences between pre- and posttraining scans were then computed using AFNI’s 3dttest, and only results that were significant at the $p < .05$ corrected level are reported.

Of the areas that showed greater functional connectivity with the functionally derived right mFG ROI during the posttraining compared with pretraining scan, mean beta coefficients were extracted for each participant to run correlations between functional connectivity strength and a behavioral measure of reading fluency.

**RESULTS**

**Behavioral Results**

Story reading rate, expressed as WPM, was the main behavioral measure of training and final reading achievement. Across the 4 days of story reading, there was a main effect of Story level, such that groups got faster across the four levels, $F(3, 72) = 4.07, \eta^2 = .15, p = .01$. However, there was no main effect of Group, $F(1, 25) = .01, \eta^2 = .00, p = .95$, and no significant interaction between Story level and Group, $F(3, 72) = 2.05, \eta^2 = .08, p = .12$. There was also no significant difference between posttraining performance on the FORT, $t(25) = .99, d = .38, p = .33$. These results indicate that once Faceabary participants acquired the Faceabary grapheme inventory and practiced single-word decoding, their text reading fluency developed at a pace comparable to that observed for the acquisition of the alphabetic FaceFont system.

**fMRI Results Posttraining: Group Differences**

**Group Comparisons in Literature-derived mFG ROI Selectivity**

A measure of the functional response to the learned artificial orthography (Face-words − Patterns) was extracted from the two mFG literature-derived ROIs in each participant. The resulting values were submitted to a $2 \times 2$ (ROI × Group) ANOVA. There was no main effect of ROI, $F(1, 25) = 1.74, \eta^2 = .06, p = .20$, but there was a main effect of Group, $F(1, 25) = 4.74, \eta^2 = .16, p = .039$, such that Faceabary participants had greater activation levels than FaceFont participants. Importantly, there was a significant interaction between ROI and Group, $F(1, 25) = 4.45, \eta^2 = .15, p = .045$, such that there were group differences, depending on the ROI (see Figure 2). Follow-up $t$ tests comparing group differences in each ROI separately revealed that there was greater selectivity in the right mFG ROI in the Faceabary as compared with the FaceFont group, $t(25) = 2.74, d = 1.11, p = .01$. However, there were no significant differences...
These results suggest that the alphabetic FaceFont group is more left-lateralized, whereas the Faceabary group is more bilateral. To confirm this interpretation, a laterality index (LI) was calculated for each participant (Right mFG ROI Selectivity − Left mFG ROI Selectivity). The Faceabary group exhibited a significantly more bilateral response than the FaceFont group, \( t(25) = 2.11, d = .83, p = .045 \) (LI_{FaceFont} = −2.74, LI_{Faceabary} = .64). To investigate this result further, the relation between individual differences in mFG laterality and reading skill (mean FORT z-score) was examined. The correlation between LI and mean FORT z-score in the Faceabary group was not significant, but it was in the positive direction, such that individuals with more left lateralized fusiform selectivity were faster readers, \( r(11) = .58, p = .048 \) (see Figure 3). To test whether the correlations differed significantly across groups, a comparison of two regression models was conducted, with LI as the dependent measure. Model 1 was a main effects model with two predictor variables, mean FORT z-score and group (FaceFont, Faceabary).

Model 2 added an additional interaction term between mean FORT z-score and group. A significant Group × Mean FORT z-score interaction term would demonstrate a different relationship between mean FORT z-score and LI for one group compared with the other. On its own, Model 1 accounted for 15% of the variance in LI, \( R^2 = .15, p = .14 \). Model 2, including the interaction term, accounted for 34% of the variance in LI, \( R^2 = .34, p = .02 \). A comparison of the two models reveals that the FORT × Group interaction term significantly improved the fit of the model, \( F(1, 23) = 6.49, p = .02 \). This suggests that the mean FORT z-score differentially predicts LI in the two groups.

**Group Comparisons Whole-brain Analysis**

A whole-brain analysis was used to investigate group differences in face-word processing that might be located outside the left and right literature-derived mFG ROIs. We conducted a 2 × 2 voxel-wise ANOVA (Group × Stimulus) to identify areas that showed a significant interaction such that there was relatively greater activity for face-words compared with patterns in the Faceabary group compared with the FaceFont group. We detected one voxel cluster in the right mFG that showed a Group × Stimulus interaction (peak voxel: 39, −62, −4, subpeak: 41, −51, −14), \( F(1, 25) = 43.29, \eta^2 = .60, p < .001 \), confirming a similar result from the literature-derived mFG ROI analysis. The significant interaction is a result of a different pattern of activation for face-words versus patterns between the two groups. For the Faceabary group, there was greater activation for face-words compared with patterns, \( t(14) = 2.35, d = .96, p = .034 \). In contrast, for the FaceFont group, there was greater activation for patterns compared with face-words, \( t(11) = −4.85, d = 1.97, p = .001 \). There were no regions that showed the opposite effect for FaceFont greater than Faceabary.

**Faceabary Group: Pre- versus Posttraining Differences**

**fMRI Differences Pre versus Post**

Both pretraining and posttraining scans were available in the Faceabary group. A 2 × 2 voxel-wise ANOVA (Session ×
Stimulus) was conducted to identify any areas that showed a training effect in the Faceabary group. A similar analysis was not possible for the FaceFont group, because these data came from a previously published study that did not include a pretraining imaging session.

Seven areas showed a significant interaction, such that there was relatively greater activation for face-words than patterns in the postscan compared with the prescan. The areas with this pattern include regions associated with phonological processing, such as the left inferior frontal gyrus (BA 44), bilateral anterior insula, and left supramarginal gyrus (see Table 1). These results suggest that Faceabary participants were engaging in phonological decoding processes even in the absence of directions to covertly read words. The opposite pattern, greater activation for face-words compared with patterns in the prescan compared with postscan, was observed in the right amygdala, and other regions shown to be part of the distributed network involved in face processing (such as STS and orbital frontal cortex; Fox, Iaria, & Barton, 2009; Ishai, 2008; Ishai, Schmidt, & Boesiger, 2005).

Functional Connectivity

Unexpectedly, even though group (FaceFont vs. Faceabary) differences were observed in the right mFG (right literature-derived mFG ROI and a functionally derived mFG ROI in the whole-brain analysis), the pre- versus posttraining analysis within the Faceabary group did not reveal a significant training effect in this brain area. One hypothesis for this null finding was that, during pretraining, the right mFG was engaged in face processing and thus showed a selectivity for faces compared with patterns, whereas after training the same area was engaged in the phonological mapping of the face-word. In this scenario, the right mFG could show similar selectivity for face-words compared with patterns, but for different reasons. To test this hypothesis, we compared functional connectivity in the pre- versus postscan, using the right functionally derived mFG ROI from the whole-brain analysis as the seed region (Table 2). We found that there was greater connectivity with this right mFG ROI in the postscan compared with prescan in several reading-related areas such as the left mFG (peak: $-49, -55, -14$), right parietal lobule (Das, Bapi, Padakannaya, & Singh, 2011; Das, Kumar, Bapi, Padakannaya, & Singh, 2009), and left angular gyrus (Pugh et al., 2000; Horwitz, Rumsey, & Donohue, 1998). Conversely, there was relatively greater functional connectivity between this right mFG ROI and surrounding cortex in the right fusiform gyrus ($32, -47, -10$) before training.

### Functional Connectivity Correlations with Behavior

Next, we examined whether any of the areas that showed greater functional connectivity with the right functionally derived mFG ROI posttraining versus pretraining were correlated with behavioral measures of reading skill (mean FORT $z$-score). The only area to show a significant

<table>
<thead>
<tr>
<th>Location</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>No. of Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post &gt; Pre</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Left</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insula</td>
<td>13</td>
<td>-33</td>
<td>30</td>
<td>-5</td>
<td>55</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>6</td>
<td>-2</td>
<td>23</td>
<td>34</td>
<td>21</td>
</tr>
<tr>
<td>Inferior frontal gyrus/precentral gyrus</td>
<td>44/6</td>
<td>-45</td>
<td>20</td>
<td>11</td>
<td>94</td>
</tr>
<tr>
<td>Inferior parietal lobule/supramarginal gyrus/angular gyrus</td>
<td>40/39</td>
<td>-39</td>
<td>-42</td>
<td>24</td>
<td>55</td>
</tr>
<tr>
<td><strong>Right</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insula</td>
<td>13</td>
<td>30</td>
<td>33</td>
<td>-2</td>
<td>33</td>
</tr>
<tr>
<td>Pre &gt; Post</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Left</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>9</td>
<td>-5</td>
<td>70</td>
<td>14</td>
<td>74</td>
</tr>
<tr>
<td>Cingulate</td>
<td>31</td>
<td>-2</td>
<td>-42</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td><strong>Right</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amygdala</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle temporal gyrus/superior temporal sulcus</td>
<td>21</td>
<td>52</td>
<td>2</td>
<td>-24</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 1. Differences for Face-words versus Patterns in Pre- or Posttraining

Hirshorn et al. 889

Downloaded from http://www.mitpressjournals.org/doi/pdfplus/10.1162/jocn_a_00940 by guest on 30 May 2021
correlation between posttraining functional connectivity strength with the right mFG ROI and reading skill was in the left mFG (peak: $-49, -55, -14$; see Figure 4). Before training, there was no correlation between the functional connectivity of the left and right mFG and reading skill (mean FORT $z$-score). After training, a significant correlation emerged, $R^2 = .31$, $F(1, 13) = 5.35, p = .04$.

**DISCUSSION**

This study used two artificial orthographies to examine the effect of mapping principle on fusiform laterality. We compared a syllable-based and an alphabetic phoneme-based writing system and found evidence of a more bilateral response in the mFG when the mapping principle of an orthographic stimulus is nonalphabetic. Training in a syllable-based system, as compared with a phoneme-based system, resulted in a more bilateral mFG response to orthographic stimuli. Furthermore, within the group trained on the syllable-based system, the degree to which the learned orthography elicited a bilateral response was correlated with the behavioral outcome of greater reading fluency. Greater bilateral mFG processing in the syllable-based system compared with the phoneme-based system in this study is consistent with studies comparing Chinese, with a larger mapping principle grain size, to alphabetic

<table>
<thead>
<tr>
<th>Location</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>No. of Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre &gt; Post</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>25</td>
<td>-23</td>
<td>-2</td>
<td>33</td>
<td>27</td>
</tr>
<tr>
<td>Right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulate/precuneus</td>
<td>23</td>
<td>-2</td>
<td>-47</td>
<td>39</td>
<td>31</td>
</tr>
<tr>
<td>Fusiform gyrus</td>
<td>37</td>
<td>32</td>
<td>-47</td>
<td>-10</td>
<td>21</td>
</tr>
<tr>
<td><strong>Post &gt; Pre</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precentral/middle frontal gyrus</td>
<td>6</td>
<td>-41</td>
<td>-5</td>
<td>57</td>
<td>32</td>
</tr>
<tr>
<td>Medial frontal gyrus</td>
<td>6</td>
<td>-5</td>
<td>-11</td>
<td>72</td>
<td>32</td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>23</td>
<td>-2</td>
<td>-23</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>20</td>
<td>-59</td>
<td>-44</td>
<td>-16</td>
<td>42</td>
</tr>
<tr>
<td>Caudate</td>
<td></td>
<td>-26</td>
<td>-44</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>5</td>
<td>-2</td>
<td>-44</td>
<td>69</td>
<td>47</td>
</tr>
<tr>
<td>Inferior temporal gyrus/fusiform gyrus</td>
<td>37</td>
<td>-49</td>
<td>-55</td>
<td>-14</td>
<td>46</td>
</tr>
<tr>
<td>Lingual gyrus</td>
<td>18</td>
<td>-1</td>
<td>-73</td>
<td>2</td>
<td>90</td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>19</td>
<td>-37</td>
<td>-76</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>Angular gyrus</td>
<td>39</td>
<td>-34</td>
<td>-76</td>
<td>32</td>
<td>175</td>
</tr>
<tr>
<td>Right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>8</td>
<td>2</td>
<td>56</td>
<td>36</td>
<td>196</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>8</td>
<td>26</td>
<td>29</td>
<td>54</td>
<td>59</td>
</tr>
<tr>
<td>Anterior cingulate gyrus</td>
<td>25</td>
<td>5</td>
<td>14</td>
<td>-7</td>
<td>75</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>6</td>
<td>38</td>
<td>-2</td>
<td>60</td>
<td>35</td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>20</td>
<td>50</td>
<td>-5</td>
<td>-37</td>
<td>22</td>
</tr>
<tr>
<td>Cerebellar tonsil</td>
<td>11</td>
<td>-41</td>
<td>-31</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Inferior/superior parietal lobule</td>
<td>40</td>
<td>44</td>
<td>-55</td>
<td>50</td>
<td>22</td>
</tr>
<tr>
<td>Precuneus/cuneus</td>
<td>19</td>
<td>5</td>
<td>-83</td>
<td>39</td>
<td>252</td>
</tr>
</tbody>
</table>
writing systems (English: Nelson et al., 2009; French: Szwed, Qiao, Jobert, Dehaene, & Cohen, 2014). Both Nelson et al. (2009) and Szwed et al. (2014) found that Chinese and alphabetic scripts elicit left mFG word selectivity, whereas only Chinese elicited right mFG word selectivity. Furthermore, an increase in right mFG functional connectivity with the left mFG was observed in the syllable-based group after training. This demonstrates that the right mFG can become engaged with the typical left-lateralized reading network after syllable-based training. These results support the notion that phonological demands vary across differences in the grain size of orthographic-phonological mapping, and these differences can lead to cross-linguistic differences in the lateralization of orthographic processing in the mFG.

As one caveat, we acknowledge that our analysis approach, which involved literature-derived ROIs based upon a VWFA coordinate drawn from the literature, did not ensure that we fully captured the VWFA (and its right hemisphere homologue) in each participant. This is because there is participant-specific variation in the location of the VWFA (Glezer & Riesenhuber, 2013). Nevertheless, the same pattern of results was observed across ROIs centered on alternative loci in the same vicinity (see Footnote 1), which suggests that the results of this study are robust and likely representative of the results that would have emerged from an analysis involving individually localized VWFA ROIs.

It is also important to note that participants did not perform an explicit task in the scanner. Consequently, potential behavioral differences between the two groups cannot be ruled out, such as differences in decoding effort or strategies. A related concern is that, because of the different number of graphs to be learned in the two behavioral training paradigms, there were necessary differences in the training protocol that could have elicited different decoding strategies. For example, the FaceFont group practiced the same words over the course of 4 days, which could have encouraged the adoption of a whole-word decoding. However, there are several pieces of evidence that make us believe that this is unlikely or would not have made a difference in our behavioral or neural outcome measures. First, even after many exposures of the same words, FaceFont participants’ word decoding did not show signs of fluency that would be expected from a sight-word reading or a whole-word decoding strategy. Although FaceFont reading latencies became faster, they were around 5000 msec on average after training (Moore, Durisko, et al., 2014), which suggests a more labored decoding strategy. Second, a whole-word reading strategy would not have been an efficient or sufficient strategy for story reading, because most words had never been seen before. Lastly, the “face-words” shown in the scanning sessions did not overlap with any of the words used in the behavioral training for either group, so it is unlikely that those particular words would have been decoded using a whole-word “sight word” strategy. Thus, we interpret group differences in mFG laterality as reflecting the use of a writing system with a phoneme-versus-syllable-based mapping principle.

Our results converge with past studies and also extend past findings. Mounting evidence supports the notion that visuoperceptual processing alone cannot explain laterality difference in mFG activation across writing systems. This further supports the proposal that

| Table 3. FORT Behavioral Measures (Mean and SD) |
|----------------|----------------|----------------|----------------|
|               | Words per Minute (WPM) | Total Number of Words | Deviations from Print (Accuracy) | Comprehension (% Correct) |
| FaceFont      | 18.37 (4.61) | 398                 | 7.96 (6.50)  | 0.87 (.08)            |
| Faccabary     | 16.6 (4.58)  | 398                 | 4.45 (5.44)  | 0.89 (.03)            |

Figure 4. Greater functional connectivity between right functionally derived mFG ROI and left fusiform posttraining. The correlation between functional connectivity between the right and left fusiform and reading rate increased after training. There was a significant positive correlation between left and right fusiform functional connectivity and reading skill (mean FORT z-score) after training such that greater connectivity was correlated with greater reading skill.
phonological mapping to orthography helps shape the selectivity for words compared with other kinds of visual stimuli in the mFG and the extent to which the right mFG also shows this selectivity (Hsiao & Lam, 2013). This proposal also provides a framework to reinterpret past work comparing different writing systems. For example, Duncan and colleagues compared two real writing systems that varied in grain size (morphosyllabic Kanji vs. syllabic Hiragana) and found greater connectivity between the left and right fusiform gyri in the system with the larger grain size of mapping (Kanji; Duncan et al., 2014). They attributed this finding to greater visual demands imposed by the more visually complex Kanji, which in turn requires greater right fusiform gyrus integration into the left language network. However, two studies have elegantly controlled for visual content (Mei et al., 2013; Yoncheva et al., 2010) while also varying whether participants received alphabetic versus logographic instruction. Both studies found greater bilateral selectivity in the fusiform gyrus for logographic-trained participants. This suggests that it is the grain size of the linguistic mapping principle, rather than the visuo-perceptual qualities of a graphemic inventory, that determines the laterality of the fusiform gyri for reading. Like others, we find that the mapping principle or unit of instruction of a writing system can change the laterality of orthographic processing within the mFG (Yoncheva et al., 2010, 2015; Mei et al., 2013). We extend this past work by showing that laterality effects can also be seen with a linguistic mapping principle smaller than that in a logographic system, but larger than a phonemic system: a syllable-based system. Unlike a logographic system, our Faceabary syllable-based system requires phonological decoding as in a phonomeme-based system; the basic units are just larger.

Although both FaceFont and Faceabary used a similar set of stimuli, there were visual differences across groups that could have affected fusiform gyrus laterality. One example is that the face graphemes in the syllable-based Faceabary had emotional content that corresponded with vowel information. The emotional content of the faces in Faceabary may have elicited greater right mFG activation, because emotional valence has been shown to modulate right amygdala and also the right-lateralized mFG face area (Ishai, 2008; Vuilleumier & Pourtois, 2007; Ishai et al., 2005; Adolphs, 2002). However, we observed opposite training patterns in the right fusiform gyrus and the right amygdala. There was greater activation in the right amygdala for faces than patterns before training, when the unlearned faces should have been processed as faces, whereas there was less activation for faces than patterns after training. This suggests that any visuo-affective differences between FaceFont and Faceabary are unlikely to have caused the observed difference in mFG laterality after training.

Additional differences between Faceabary and FaceFont stimuli could be that more attention was required of fine facial detail in the Faceabary group. However, one neuro-imaging study revealed a right fusiform advantage for the processing of faces as a whole and a left fusiform superiority for the processing of facial features (Rossion et al., 2000). Similarly, another recent study reported that successful episodic memory encoding of faces relied on the left fusiform cortex because of the involvement of feature/part information processing (Mei et al., 2010). Thus, if greater featural facial processing demands drove group differences, we may have expected the opposite result of greater left fusiform selectivity for Faceabary. The fact that we observed greater selectivity in the right fusiform suggests that the larger mapping principle grain size is more plausible that differences in attention to visual features.

Past work also speaks to the converse hypothesis that facial expressions in Faceabary would require more holistic configurational processing. For example, Moore, Durisko, et al. (2014) kept the decoding training constant across two alphabetic writing systems that differed in visual processing demands (faces vs. “KoreanFont” line segments). They reasoned that if holistic visual processing drives right mFG engagement, greater right mFG activation should be observed for the FaceFont. They observed similar left hemisphere lateralization for both writing systems, which led them to rule out visual processing demands as a major determinant of VWFA lateralization. Mei et al. (2013) and Yoncheva et al. (2010) used the same exact visual stimuli, but varied the decoding training (phoneme vs. whole word), and found greater right mFG engagement for the whole-word trained participants. Taken together, these results have influenced our interpretations of the current results to reflect group differences caused by differences in linguistic mapping principles as opposed to differences in holistic visual processing.

As further evidence, the increase in functional connectivity between the right and left mFG after training demonstrates how the functional underpinnings of the engagement of right mFG can change. It also offers insight into why we observed group differences in the selectivity of the right mFG. Whereas the right mFG ROI defined from group differences, showed greater connectivity with the neighboring right mFG region before training, when the face–syllable mappings had not yet been learned, it showed greater connectivity with the left mFG after training, when the linguistic mappings had been learned. Finally, the fact that individual differences in right mFG activation and in functional connectivity to the left mFG correlated with a reading outcome measure clearly supports the idea that the right mFG became involved in Faceabary reading.

Although the focus of the current article was to disentangle two potential hypotheses (e.g., visual spatial processing demands vs. mapping principle) regarding cross-linguistic differences in mFG laterality, the results also contribute to an ongoing debate about what determines the location and function of the VWFA within the left mFG. One hypothesis suggests that the VWFA is fundamentally sensitive to visual statistics and its location results from neuronal recycling of inferotemporal cortex used for the detection of shapes in foveal vision (Dehaene & Cohen, 2007). In contrast, a
competing hypothesis proposes that the VWFA is sensitive to linguistic and orthographic properties of words (Price & Devlin, 2011) and is in its location due to connectivity between the visual and linguistic systems within the mFG. We suggest that these two proposals do not need to be mutually exclusive. Certainly any neural functioning related to reading would have to result from some form of neuronal recycling, as reading is a relatively new invention in the timeframe of evolution. Interestingly, there are laterality differences in the acoustic processing of speech such that phonemes are left-lateralized, whereas syllable level acoustics (measured as theta-oscillations) are right-lateralized (Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012; Poeppel, Iksandi, & van Wassenhove, 2008; Giraud et al., 2007; Hickok & Poeppel, 2007). We therefore propose that neuronal recycling is driven by the binding of visual and speech-based information, with the degree of laterality determined by the units of speech associated with the decoding of the orthographic forms.

One last important component of the current study is that participants in both groups learned to read a productive orthography at the story level. The development of a syllable-based orthographic system for English may have practical benefits. Past research in naturally occurring and artificial orthographies suggests the idea that learning a syllable-based writing system can elicit improvements in pronunciation (e.g., “pa + at = pat” is more effective than “p + a + t = pat”), has also been observed in alexia and dyslexic individuals (Bowes & Martin, 2007; Friedman & Lott, 2002). Further support comes from a patient with pure alexia and left mFG damage who had impaired phoneme-grapheme mapping but relatively preserved syllable-grapheme mapping using artificial orthographies (Moore, Brendel, et al., 2014). An interesting possibility to explore in future research rests upon the logic that using a syllable-based orthography may encourage the use of decoding strategies that rely to a lesser extent on left hemisphere reading areas. This could be especially relevant to dyslexic readers who have been shown to have less left-lateralized reading networks (Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Shaywitz et al., 2002). Thus, future research could test whether dyslexic readers may benefit to a greater extent from a syllable-based orthography than typical readers and open a new area for investigating remediation strategies for at-risk readers.

Reprint requests should be sent to Elizabeth A. Hirshorn, Learning Research and Development Center, 3939 O’Hara Street, Pittsburgh, PA 15260, or via e-mail: hirschorn@pitt.edu.

Note

1. The same pattern of results was observed using ROIs derived from a “quasiloocator” defined as KoreanFont versus Patterns (an approach successfully deployed in our previous work; Moore, Durisko, et al., 2014) as well as ROIs centered on alternative VWFA coordinates reported in published work (McCamliess et al., 2003; Cohen et al., 2002; Dehaene et al., 2001).

REFERENCES


