



Published in final edited form as:

J Cogn Neurosci. 2014 April ; 26(4): 896–913. doi:10.1162/jocn_a_00506.

Learning to Read an Alphabet of Human Faces Produces Left-lateralized Training Effects in the Fusiform Gyrus

Michelle W. Moore^{1,2}, Corrine Durisko³, Charles A. Perfetti^{1,3,4}, and Julie A. Fiez^{1,3,4}

¹University of Pittsburgh, Pittsburgh, PA

²West Virginia University, Pittsburgh, PA

³Learning Research and Development Center, Pittsburgh, PA

⁴Center for the Neural Basis of Cognition, Pittsburgh, PA

Abstract

Numerous functional neuroimaging studies have shown that most orthographic stimuli, such as printed English words, produce a left-lateralized response within the fusiform gyrus (FG) at a characteristic location termed the visual word form area (VWFA). We developed an experimental alphabet (FaceFont) comprising 35 face–phoneme pairs to disentangle phonological and perceptual influences on the lateralization of orthographic processing within the FG. Using functional imaging, we found that a region in the vicinity of the VWFA responded to FaceFont words more strongly in trained versus untrained participants, whereas no differences were observed in the right FG. The trained response magnitudes in the left FG region correlated with behavioral reading performance, providing strong evidence that the neural tissue recruited by training supported the newly acquired reading skill. These results indicate that the left lateralization of the orthographic processing is not restricted to stimuli with particular visual-perceptual features. Instead, lateralization may occur because the anatomical projections in the vicinity of the VWFA provide a unique interconnection between the visual system and left-lateralized language areas involved in the representation of speech.

INTRODUCTION

Cognitive neuroscience has played an important role in furthering our understanding of the cognitive processes and neural substrates that support skilled reading. An area within the left fusiform gyrus (FG), termed the visual word form area (VWFA), has been a particular focus of research. This region exhibits a relatively selective response to orthographic stimuli, as compared with other categories of visual objects (McCandliss, Cohen, & Dehaene, 2003). Within the domain of orthography, the VWFA is most robustly activated by words and pronounceable nonwords, with lesser responses usually seen for consonant letter strings, stimuli composed of letter-like forms, or grapheme strings from an unfamiliar language (Polk, Stallcup, Aguirre, & Alsop, 2002; Cohen et al., 2000). These properties suggest that

the VWFA is a core—perhaps even obligatory—part of the pathway by which the perceptual analysis of written words ultimately provides fluent (fast and accurate) access to representations of phonology and meaning.

There are competing views on the underlying function of the VWFA. Points of disagreement include whether the region is unimodal (Price & Devlin, 2003), whether it is uniquely activated by orthographic stimuli (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2007), and where it lies on the processing stream that supports orthographic-to-phonological transformation (Vinckier et al., 2007; Pugh et al., 2000). This study homes in on a question that has received less attention in the literature: Why is orthographic processing typically lateralized to the left FG? We evaluate two accounts for the lateralization of orthographic processing, both of which arise from a consideration of the unique properties of orthographic systems and which are inspired by similar distinctions outlined in the speech (Shtyrov, Pihko, & Pulvermuller, 2005; Melamed & Zaidel, 1993) and word recognition literatures (Ellis, Ansorge, & Lavidor, 2007; Marsolek & Deason, 2007).

Printed words may tend to invoke left-lateralized activation within the FG because the development of writing systems has led to mappings between visual graphic units and linguistic knowledge (e.g., speech segments) acquired through spoken language. The sector of the FG that has been associated with the VWFA may be special because it is uniquely positioned, by virtue of its anatomical connections, to serve as a bridge to left-lateralized language areas (Reinke, Fernandes, Schwindt, O'Craven, & Grady, 2008; Vigneau, Jobard, Mazoyer, & Tzourio-Mazoyer, 2005). Because automatic word identification is a central feature of skilled reading (Perfetti & Hart, 2002), a logical extension is that the VWFA may thus play a central role by acting as a bridge between visual perception and pre-existing linguistic knowledge.

This linguistic bridge account of the VWFA can provide a compelling explanation for the syndrome of pure alexia, in which reading occurs through the use of a very slow and effortful “letter-by-letter” strategy (Montant & Behrmann, 2000; Black & Behrmann, 1994), although higher-level language functions are intact. This syndrome has been associated with damage to the VWFA (Gaillard et al., 2006; Cohen et al., 2003); the permanent loss of a critical bridge area into the language system could explain why efforts at remediation have achieved only modest success (Seki et al., 2001; Maher, Clayton, Barrett, Schober-Peterson, & Gonzalez-Rothi, 1998; Behrmann & McLeod, 1995; Arguin & Bub, 1994; Daniel, Bolter, & Long, 1992; Behrmann, Black, & Bub, 1990).

Printed words, alternatively, may tend to invoke left-lateralized activation within the FG, not because the items are linguistic in nature but because printed words impose a high demand on a particular type of perceptual analysis (McCandliss et al., 2003). Orthographic systems were specifically developed to be reproducible by writing, with a concomitant emphasis on efficiency of production and robustness to variability in form (Changizi & Shimojo, 2005). As a consequence, orthographic systems may naturally promote the use of feature-based visual analysis strategies, in which characteristic differences in the lengths, orientations, and junctions of line segments are used to discriminate between component graphemes (e.g., the letters in the Roman alphabet) and to extract graphemic prototypes. Furthermore, the high

degree of perceptual similarity across graphemes may necessitate the extraction of high-spatial frequencies. The left hemisphere has been particularly implicated in feature-based, high-spatial frequency analysis (Woodhead, Wise, Sereno, & Leech, 2011; Mecacci, 1993).

Natural orthographies typically consist of letters or characters that are composed of line segments so that feature-based, high spatial frequency analysis is paired with the linguistic processes involved during reading. Accordingly, it is difficult to distinguish the two accounts for the characteristic left-lateralization of orthographic processing on the basis of natural orthographies. Contrasting orthographies that have an identical linguistic function but differ substantially in their visual perceptual characteristics would allow a test of these accounts.

This line of reasoning led us to develop two experimental alphabets to represent the speech sounds of English. Our novel “FaceFont” alphabet uses faces as graphemes (Tottenham et al., 2009), and our “KoreanFont” alphabet uses Korean characters as graphemes (Figure 1; see also Appendix A). We chose faces for two reasons. First, prior work has demonstrated that face processing can induce a right-lateralized response in the FG at a region that has been termed the “fusiform face area” (FFA; Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher, McDermott, & Chun, 1997). Second, faces are perceptually unlike prototypical letters, and compared with printed words, their processing appears to be more reliant on holistic and low-frequency visual analysis mechanisms (Goffaux & Rossion, 2006). Thus, their use as letters creates an orthography that provides a highly distinctive perceptual contrast to the forms of natural orthographies. We chose Korean letters (the Hangeul alphabet) as the basis for our comparison because they consist of the visual components, for example, line segments, that are typical of all alphabets but are unknown to our native English speakers. Learned responses to FaceFont stimuli in the left, but not the right, FG would indicate that pairing faces with linguistic information produces the characteristic lateralization of orthographic processing as predicted by the linguistic bridge account. Learned responses to FaceFont stimuli in the right FG would favor a visuo-perceptual account for the lateralization of orthographic processing.

METHODS

Participants

Twenty-four participants (11 men) completed the 2-week training portion of this study (mean age = 21.1 years, $SD = 1.8$ years). Training study participants were recruited through fliers posted around the University of Pittsburgh campus. Twelve additional controls (5 men) completed a single neuroimaging session (mean age = 22.3 years, $SD = 2.7$ years). These controls were recruited from a database of participants who had participated in previous neuroimaging studies and had indicated that they would be interested in participating in future studies.

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 fMRI 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.

Procedure

Participants selected for the training portion of the study were pseudorandomly assigned to train on either the FaceFont stimuli or the KoreanFont stimuli. FaceFont-trained (FF; $n = 12$) and KoreanFont-trained (KF; $n = 12$) participants completed 1- to 2-hr training sessions consisting of three components: phoneme training (Day 1), word level training (Days 2–5), and story level training (Days 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 behavioral testing and fMRI 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 and KoreanFont reading (see Table 1 for training schedule). Visual items throughout all sessions were presented exclusively in the training fonts and were paired with spoken English equivalents at the beginning of training, as necessary. No printed English was used except for the initial basic instructions of a task.

Alphabets—The mapping principles that were used to associate our experimental graphemes to English speech sounds were equivalent for FaceFont and KoreanFont. Using a consistent alphabetic system with a one-to-one grapheme-to-phoneme correspondence, we used 35 letter–sound pairs to represent all sounds in English (Figure 1). There were five exceptions in which a grapheme (i.e., a face or a Korean character) represented two similar sounds.

Phoneme Training—In Session 1, participants completed phoneme training. Using the E-prime computer program for psychological experiments (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. Only one person (assigned to the KF group) could not advance to the word level training because they could not meet the 90% criterion within the first

session. All participants who advanced to word level training 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/, /i/, /f/. Using a keypress, participants could play any of the three phonemes individually or could play the entire word, if necessary.

Story Level Training—During each session in the second week of training (Sessions 6–9), participants 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 1 stories and progressing one level each session to longer and more complex stories (up to Level 4). Story level reading performance was measured by the number of words read per minute (WPM) for each story.

Word Tests—After completing each word level and story level training session (Sessions 2–9), participants took single-word reading tests on the computer. Participants were presented with 45 items one at a time and asked to read each one aloud as quickly and as accurately as possible. Of the 45 items, there were 15 “old” words (words that were in the word training), 15 “new” words (words that were not in the word training task), and 15 nonwords. The 45 items were administered in random order.

To obtain purer accuracy and latency measures on real word reading without the influence of nonwords intermixed with real words, during the second week of training we followed each word test with a supplemental test. The supplemental tests were brief, consisting of 12 real words. The real words were a selection of “new” word items from Word Test 1 (WT1) through WT4.

Typically, word tests were administered after the participants completed a training session. One exception to this occurred during the first session of the second training week (Session 6), in which two sets of word tests were administered—WT5 and Supplemental Word Test 5 (WT5S) were administered at the start of the session, and WT6 and WT6S were administered at the end of the session. We administered WT5/WT5S at the start of the session to measure retention of the training font after two weekend days without training. Because WT5 was administered under different conditions than the other word tests (preceding training rather than following training), it was not included in the primary analyses.

Reading latencies were measured from word onset to when the participant elicited the next item with a mouse click. This diverges from the more conventional latency measure from word onset to the participant’s voice onset when beginning to read the word; however, because there was such variability across participants in decoding words (some participants

sounded out words silently, whereas some did so aloud), the latency measure used in this study provided a more consistent measure across participants. Only reading latencies from items read correctly were considered in analyses. On a few trials, participants changed their answer, asked the experimenter a question, or had some interruption while a test item was on the screen. The response times for these items were excluded from analyses.

Standardized Reading Test—During the final session (Session 10), 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. The stories were administered and scored according to the standardized test protocol. 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 acoustically). Raw scores for the following were obtained: accuracy, time, fluency (accuracy + time), and comprehension. Standardized scores were not used because (1) the stories had been transcribed into our experimental fonts and (2) participants in our study exceeded the maximum age on which normative data were collected for this test. In addition to the raw scores used from the standardized test, we also computed WPM for each story.

fMRI Data Acquisition—A 3-T head only Siemens Allegra magnet and standard radio frequency coil were used for all MR scanning sessions at the University of Pittsburgh. Prior to functional scanning, structural images were collected using a standard T2-weighted pulse sequence in 38 contiguous slices ($3.125 \times 3.125 \times 3.2$ mm voxels) parallel to the AC-PC plane. Functional images were collected in the same plane as the anatomical series using a one-shot EPI pulse sequence (repetition time = 2000 msec, echo time = 25 msec, field of view = 200 mm, flip angle = 70°).

Training study participants were asked to passively view three types of stimuli—FaceFont words, KoreanFont words, and patterns (used as a baseline). They completed two functional runs designed identically. Each run contained twenty-one 20-sec epochs, seven epochs of each stimulus type presented in random order. Each epoch consisted of ten 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 words in their specific training font. Both the FF group and the KF 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 400-word training set or the word tests from the behavioral sessions. Words were presented in random order across both runs. There were 16 different patterns that were used repeatedly in random order throughout both runs (Figure 2).

The fMRI data acquisition for the control participants was identical to the data acquisition described for the trained participants. The first two functional runs for the control

participants were also identical to the trained participants' runs. Additionally, the controls completed a third run at the end of the session involving other visual stimuli that will not be discussed here.

fMRI Data Analysis—A series of preprocessing steps were conducted prior to data analysis using an integrative software package, Fiswidgets (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, Mazziotta, & 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 three-dimensional Gaussian filter (8 mm FWHM) to account for anatomical differences between participants. Images were finally converted into Talairach space (Talairach & Tournoux, 1988). Specific cerebellar lobules were designated using a cerebellar atlas (Schmahmann, Doyon, Toga, Petrides, & Evans, 2000).

Several fMRI analyses were completed to address different aspects of the data. As our principal analysis, we used the NeuroImaging Software package (NIS 3.6) to compute a whole brain, voxel-wise 2×2 mixed-model ANOVA with Training Group (FF, KF) as a between-subject factor, Font Condition as a within-subject factor (FaceFont, KoreanFont), and Subject as a random factor. The dependent measure was the mean normalized voxel signal intensity for each font condition for each participant. Because there were no fixation or rest intervals between the stimulus condition epochs, the first 8 sec of each 20-sec epoch were not included in the computation of the mean signal intensity values; this was done to minimize cross-epoch contamination of the BOLD signal. A voxel-wise significance threshold of $p < .000005$ (corrected to $p = .01$ using AFNI 3dClustSim; Cox, 1996) and contiguity threshold of five voxels was used to identify voxel clusters that exhibited significant effects of font condition, training group, and the interaction between these two factors. The interaction between font condition and training group was of particular interest, because it should reveal voxel clusters in which the responses to our stimuli were modulated by the training group assignment (i.e., it should identify the locus of training effects). The visuo-perceptual account predicts that a significant interaction should be found in both the left and the right FG: KoreanFont should elicit stronger left FG activation in KF-trained versus FF-trained participants, and FaceFont should elicit a weak left FG activation of similar magnitude in both KF-trained and FF-trained participants; conversely, FaceFont should elicit stronger right FG activation in FF-trained versus KF-trained participants, and KoreanFont should elicit a weak right FG activation of similar magnitude in both FF-trained and KF-trained participants. The linguistic bridge account, on the other hand, predicts a significant interaction of these factors in only the left FG, with a crossover pattern: the response to KoreanFont should be higher in KF-trained versus FF-trained participants,

whereas the response to FaceFont should be higher in FF-trained versus KF-trained participants.

We extracted the hemodynamic time course data from significant voxel clusters identified by our primary voxel-wise ANOVA model, with a specific focus on voxel clusters in the left or right FG (the a priori brain areas of interest). This was done to ascertain the underlying pattern of simple effects by examining differences in the percent signal change in each condition (relative to the baseline pattern condition) as a function of training group and to draw secondary comparisons between our training group participants and our control participants. As a complementary approach for examining the pattern of simple effects, we performed a secondary conjunction analysis. This conjunction analysis compared the voxel-wise statistical contrast map for the response to FaceFont in trained versus untrained (KF-trained and control) participants and the map for the response to KoreanFont in trained versus untrained (FF-trained and control) participants.

Additional secondary analyses targeted the identification of FG voxels with differential responses to face and word-like stimuli, that is, voxels that should fall at or near the FFA and its potential left-hemisphere homologue (l-FFA) and voxel clusters that should fall at or near the VWFA and its potential right-hemisphere homologue. One secondary analysis was performed because the primary analysis did not reveal a significant interaction in the right FG, leading us to seek further confirmation of this null result. Accordingly, we conducted a group level, voxel-wise analysis that contrasted FaceFont versus patterns, with all of the participants from the three experimental groups included. A voxel-wise significance threshold of $p < .000005$ (corrected to $p = .01$ using AFNI 3dClustSim; Cox, 1996) and contiguity threshold of five voxels was used to identify voxel clusters that exhibited significant effects of stimulus condition. We then searched the resulting statistical map to identify voxel clusters in the left and right FG, extracting the time course data from each identified voxel cluster. The extracted data for each voxel cluster were submitted to a 3×2 mixed-design ANOVA with Training Group (FF, KF, controls) as a between-subject factor and Font Condition (FaceFont, KoreanFont, as percent signal change relative to the pattern baseline) as a within-subject factor. Simple effects analyses were also computed to draw comparisons between trained and untrained participants for each font type.

Another secondary analysis was performed because the primary analysis did reveal significant training effects in the left FG for both KoreanFont and FaceFont. This led us to investigate the correspondence between these training effects and the VWFA. We chose not to expose our training participants to printed English stimuli during training or during our scanning sessions because we wanted to discourage any association between our artificial orthographies and printed English words. As a consequence, we could not “localize” the VWFA via a contrast between printed English words versus nonorthographic stimuli, as is commonly done in the literature (see Appendix B). However, all of the participants in our experiment were exposed to the KoreanFont stimuli, and there is evidence that letter-like stimuli, even for an unknown orthography, can produce activation within the VWFA (e.g., Ben-Shachar et al., 2007). This led us to reason that the KoreanFont versus pattern contrast could serve as a proxy for a more typical VWFA functional localizer. Using this logic, we

conducted a “quasilocalizer” analysis to evaluate the probable impact of our training on the functional VWFA responses of our participants.

In each participant, we searched within a 12-mm spherical volume for the voxel exhibiting the largest increase in mean signal intensity for KoreanFont versus pattern epochs. The search volume was centered on the peak of the left FG voxel cluster that exhibited an interaction in our primary analysis. For each participant, we computed the mean percent signal change in response to FaceFont and KoreanFont (using patterns as a baseline) in the localized KoreanFont voxel from that participant. The signal change values were submitted to the same ANOVA model and simple effects analyses that were used for our primary analysis. It is important to note that this secondary analysis was deliberately biased toward the selection of voxels that respond to KoreanFont, because the KoreanFont > pattern contrast was used as the proxy localizer for the VWFA. Thus, stronger training effects for KoreanFont versus FaceFont would not be unexpected. The key question, however, was whether a significant FaceFont training effect could nonetheless be detected in the localized KoreanFont voxels.

A limitation of both the voxel-wise and quasilocalizer analyses is that the results are based on a comparison of response magnitudes at particular voxels that fall within the general vicinity of the VWFA. Neither addresses the question of whether the spatial topography of our training effects varies according to the trained font. This is an important question, because recent findings indicate that responses to face stimuli can be detected within the vicinity of the VWFA. The locus and interpretation of these face responses remain a point of active debate and investigation. For instance, some have found that face responses within the left FG tend to be located a few millimeters antero-medial to the VWFA and have suggested that these responses may arise from a l-FFA (Hasson, Levy, Behrmann, Hendler, & Malach, 2002). Others have identified graded sensitivity to faces within and near the VWFA and have suggested that the acquisition of literacy may “perceptually tune” the response properties of the VWFA and its surrounding neighborhood (Dehaene et al., 2010).

In a final secondary analysis, we extended our quasi-localizer analysis by also identifying, within each participant, the voxel that exhibited the largest mean increase during FaceFont versus pattern epochs. This yielded a set of localized FaceFont voxels. The x , y , and z coordinates of each localized voxel were used as the dependent measure in a mixed-effects MANOVA, with Training Group (FF, KF) as a between-subject measure and Voxel Type (FF-localized, KF-localized) as a within-subject measure. Follow-up simple contrasts and additional post hoc comparisons to the localized voxel responses from the control group were used to further investigate the pattern of effects.

If faces elicit an average response that tends to peak at a different location within the left FG than the response elicited by word-like stimuli, then a main effect of Voxel Type would be expected (i.e., the mean location of the FF-localized voxels should differ from the mean location of the KF-localized voxels). If the location of the peak response to a particular stimulus tends to be affected by training (e.g., training with FaceFont shifts the response to FaceFont stimuli toward the VWFA), then an interaction between Voxel Type and Training Group would be expected.

RESULTS

Can Participants Learn the Novel Orthographies?

Individuals learned to read their assigned orthography, and their performance generally became less variable as they progressed in training (Figure 3).¹ A 2×4 mixed-design ANOVA (Group \times Word Test, with corrected degrees of freedom for violations of sphericity using Huynh–Feldt estimates) showed that there was a significant decrease in latencies on the single-word reading tests across training sessions during the initial phase of training (Week 1 word level training), indicating that learning occurred, $F(1.93, 42.40) = 29.03, p < .001$, main effect of Word Test. The KF group was nonsignificantly faster than the FF group in reading latencies on the daily word tests ($p = .17$, main effect of Group). The two groups showed similar rates of improvement ($p = .50$, Group \times Word Test interaction). Thus, although the KF group tended to show better overall performance (at a level of low reliability) than the FF group, the learning of the new orthographies proceeded at similar rates for the two groups.

A separate 2×4 ANOVA (Group \times Word Test) was used to examine reading performance during the story level phase of training (Week 2). Latencies on the single-word reading tests further decreased across training sessions, indicating that learning continued to occur, $F(3, 66) = 11.21, p < .001$. The KF group again showed shorter reading latencies than the FF group, now more reliable statistically, $F(1, 22) = 9.26, p = .01$. However, the two groups continued to demonstrate similar rates of improvement ($p = .85$).

Thus, across the 2 weeks, word reading was generally slower for the FF group than the KF group, but the two groups showed a similar learning pace. The relatively slower reading latencies for the FF group may have been harder to detect in the initial word tests because of a higher degree of performance variability as participants learned the task, adjusted to reading at a phoneme level, and overcame memory constraints. As the impact of these initial barriers diminished, any perceptual influence of the stimuli on reading performance should become clearer.

Participants' performance while reading stories provides additional information about the learning of the two training fonts. We measured the number of WPM in the early reader stories as well as the standardized reading test (Figure 3C, D). Consistent with the daily word test results, the FF group's reading rate was slower than the KF group's rate. Notably, however, both groups were similarly sensitive to the fluctuations in text style and complexity, suggesting that the overall speed differences arose from nonlinguistic, visual factors (FF vs. KF WPM on 40 early reader stories: Pearson's $r = .70, p < .001$; FF vs. KF WPM on six stories from the standardized reading test: Pearson's $r = .85, p = .03$).

¹In the main text, we focus on the reading latencies for the word tests during behavioral training. There were no significant group or test differences in accuracy on the word tests during the behavioral training, nor was there a Group \times Test interaction in accuracy. Average accuracy for trained participants across all word tests was 91%. The brief supplemental tests (WT5S–WT9S) gave a similar pattern of results.

Investigation of Neural Responses Using fMRI

Localization of Training Effects—During the fMRI session, FF participants ($n = 12$), KF participants ($n = 11$), and untrained controls ($n = 12$) were presented with patterns (baseline), FaceFont words, and KoreanFont words using a blocked design. To probe for responsive neural regions in trained participants, as our primary analysis we computed 2×2 voxel-wise ANOVAs (Training Group \times Font). There were no significant voxel clusters that exhibited a main effect of Training Group. There were, however, significant voxel clusters in left and right occipitotemporal cortex that showed a differential response to the artificial fonts (Table 2). Such differences were expected, because it is well established that visual exposure to alphabetic characters preferentially produces activation within the left FG at or near the VWFA, whereas visual presentation of faces preferentially produces activation within the right FG at or near the FFA (Kanwisher & Yovel, 2006; Cohen et al., 2002).

Our primary objective was to identify neural regions within the left or right FG that exhibited a significant interaction pattern in which the response varied across groups based on the assigned training font. There was a significant interaction in the left FG at a location consistent with the VWFA (peak voxel: $-40, -50, -11$; cluster size: 6 voxels; see Figure 4A). Notably, no significant interaction pattern was observed in the right FG. Significant interactions were found in regions outside of the FG, such as the left inferior frontal gyrus. The locus and interpretation of these training effects will be reported in a separate paper.

We used follow-up analyses to further examine the voxel cluster within the left FG (LFG-VC), which was located in the vicinity of the VWFA, that exhibited an interaction pattern in our primary analysis (Figure 4B). As stated previously, the visuo-perceptual account predicts a stronger response in the LFG-VC for KoreanFont in trained versus untrained participants, with no response differences for FaceFont between trained and untrained participants. In contrast, the linguistic bridge account predicts a training effect for both fonts within the LFG-VC. As expected based on both accounts, the response to KoreanFont in the LFG-VC was enhanced by training, with a significantly larger response to KoreanFont words (compared with a string of pattern baseline stimuli) observed in the KF group compared with participants who did not complete KoreanFont training (KF group vs. FF group: $t(21) = 7.71, p < .001$; KF group vs. untrained controls: $t(21) = 7.04, p < .001$). The magnitude of the response to FaceFont words was also significantly greater in the FF group than the FaceFont response observed in the untrained groups (FF group vs. KF group: $t(16.81) = 3.55, p = .003$, t test with degrees of freedom corrected for violations of Levene's equality of variances assumption; FF group vs. untrained control group: $t(22) = -2.08, p = .05$). The sensitivity to training observed within the LFG-VC for both KoreanFont and FaceFont words provides compelling support for the linguistic bridge account of orthographic lateralization within the FG.

Additional support for these results was found in our secondary conjunction analysis in which we compared voxel-wise statistical contrast maps for the FaceFont response in trained versus untrained participants and the KoreanFont response in trained versus untrained participants. Results converge with those from the primary analysis: Training effects were found for both fonts within the left FG, and the identified voxel clusters partially overlapped

at a locus at or near the typical location of the VWFA. There was an absence of training effects in the right FG for either font (Figure 4C).

To provide confirmatory evidence for the null result in our primary analysis in which there was no significant Training Group \times Font interaction in the right FG, we computed a group level, voxel-wise ANOVA that contrasted FaceFont versus patterns across all three experimental groups. This analysis identified two significant voxel clusters: one located in the right FG with a location consistent with the FFA (FFA-VC; peak voxel: 35, -38, -21; cluster size: 119 voxels) and one located in a homologous portion of the left FG (l-FFA-VC; peak voxel: -40, -48, -21; cluster size: 62 voxels). The time course data from these voxel clusters were submitted to a 2×3 mixed-design ANOVA (Font \times Training Group). In the FFA-VC, there was an expected main effect of Font (FaceFont > KoreanFont as compared with pattern baseline stimuli; $F(1, 32) = 70.13, p < .001$), but no significant interaction between Training Group and Font ($p = .260$) or main effect of group ($p = .941$). Simple effects comparison between the FF-trained versus untrained groups and the KF-trained versus untrained groups were also non-significant (FF-trained vs. untrained: $p = .763$; KF-trained vs. untrained: $p = .204$). These results are consistent with the null findings in our primary analysis and provide additional support for the linguistic bridge account.

The analyses computed for the l-FFA-VC provide even further support for the results from our primary analyses. There was a significant interaction between Training Group and Font, $F(2, 32) = 8.01, p = .002$. There was also a main effect of Font (FaceFont > KoreanFont as compared with patterned baseline stimuli; $F(1, 32) = 42.59, p < .001$) and no main effect of Group ($p = .723$). The simple effects comparison between the FF-trained versus untrained groups approached significance ($p = .053$), and the KF-trained versus untrained group comparison reached significance ($p = .003$).

As an additional point, it should be noted from our primary analysis that, although trained participants from both the KF and FF groups had stronger activation patterns within the LFG-VC compared with untrained participants when viewing their training font, KoreanFont elicits a larger response magnitude than FaceFont for both trained and untrained participants. For example, all three groups had a significant response to KoreanFont words (untrained controls: $t(11) = 3.18, p = .01$; FF group: $t(11) = 3.87, p = .003$; KF group: $t(10) = 15.47, p < .001$), whereas FaceFont elicited a slightly negative response in the LFG-VC in both of the untrained groups (KF group, control group). These magnitude differences in untrained participants are consistent with prior results showing the VWFA is biased toward stimuli that have the features of a typical orthography (Ben-Shachar et al., 2007) and that preferential responses to faces tend to localize to the right FG (i.e., to the FFA) with less activation in the left FG (Cohen et al., 2002). Furthermore, this pattern of results suggests to us that the perceptual features of KoreanFont may give it a “head start” that facilitates its acquisition as a novel orthography.

These differences in the magnitude of the left FG response to KoreanFont versus FaceFont converge with the differences in reading rates of FF-trained versus KF-trained participants to indicate that the perceptual form of our orthographies did exert an influence. To evaluate this issue further, we examined the relationship between an individual’s behavioral

performance (final word test latency) and their neural response (LFG-VC % stimulus response to participant-specific training font) across both training groups. Faster reading times (shorter latencies) correlated with greater activation in the LFG-VC for the participant-specific training font (Pearson's $r = -.70, p < .001$). Thus, although the left lateralization of orthographic processing seems to be driven by the linguistic characteristics of the experimental fonts, stimulus-type training differences can be detected in the left FG and the corresponding differences in reading fluency; therefore, the perceptual qualities of a font do seem to have an impact on learnability and neural reorganization.

Relationship between Training Effects and the VWFA—The primary objective of this study was to disentangle linguistic and perceptual influences on the lateralization of orthographic processing. Nonetheless, we recognize the results naturally give rise to questions about the relationship between the locus of FaceFont training effects and the VWFA. Because the imaging design did not include a contrast between English words and a nonorthographic control condition, it cannot be fully determined whether FaceFont training modulates the response properties of the same neural tissue involved in the processing of natural orthographic stimuli (i.e., the VWFA). However, two supplemental analyses do provide provocative evidence in support of this conclusion.

For one analysis, we used the KoreanFont versus pattern contrast as a “stand-in” for a more typical VWFA functional localizer. The mean of the voxel loci found within each group of participants was comparable with those observed in our review of the recent literature, giving us confidence in this quasilocalizer approach (FF group: $-40, -49, -12$; KF group: $-42, -53, -12$; control group: $-40, -51, -13$; see Appendix B for comparisons with recent literature). We computed the mean percent signal change in response to FaceFont and KoreanFont (using patterns as a baseline) in the KF-localized voxel from each participant and then submitted these values to the same set of analyses that were applied to the signal intensity values extracted from the LFG-ROI. The results provide an even more robust demonstration of the training effect patterns observed within the LFG-ROI (Figure 5). For instance, although all of the voxels for this analysis were identified from the KoreanFont versus pattern contrast, in FF participants the voxels exhibited a greater response to FaceFont than KoreanFont. On the other hand, for KF participants, the response to KoreanFont is increased relative to that observed in the FF and control participants, and the response to FaceFont is decreased.

In a second analysis, we investigated whether the mean location of the peak response to FaceFont versus KoreanFont varies within the left FG and whether the peak response locations are affected by training. In each participant, we used a FaceFont and a KoreanFont localizer contrast to determine the location of the maximum response to each font type within a 12-mm sphere centered on the LFG-ROI from our primary analysis. The identified voxel coordinates were then used as the dependent values in a mixed-effects MANOVA. We observed a main effect of Voxel Type, $F(3, 19) = 16.70, p < .001$, which reflects the fact that FF-localized voxels tended to lie anteromedial and ventral to KF-localized voxels (Figure 6). This main effect of Voxel Type is consistent with some prior reports showing that faces tend to produce activation in both the left and right FG that lies anteromedial and ventral to that produced by words (Hasson et al., 2002).

Importantly, the main effect of Voxel Type interacted with Training Group, $F(3, 19) = 8.54$, $p = .001$. In the FF group, FF-localized voxels were shifted toward the typical location of KF-localized voxels. Furthermore, within the FF group, the FF-localized voxels did not differ in mean location from the KF-localized voxels, whereas the two localizers produced voxel sets that trended toward different mean locations in the KF and control groups. In fact, in 58% (7/12) of the participants in the FF group, the same voxel was identified by the FaceFont localizer and the KoreanFont localizer; this was true for 0% (0/11) and 8% (1/12) of the participants in the KF and control groups, respectively. In summary, these results indicate that FaceFont training shifts the response to FaceFont toward the prototypical locus of the VWFA.

Training-related shifts were also observed for the KF group. In the KF group, KF-localized voxels were shifted lateroposteriorly to KF-localized voxels in the FF and control groups and further away from the FF-localized voxels in each group. In a sense, KoreanFont training results in a response to KoreanFont that is more “VWFA-like” than usual.

One way to account for these training-related shifts is to propose that, in untrained participants, individual voxels within the left FG exhibit a noisy activation bias toward either FaceFont or KoreanFont and a noisy tendency toward topographic organization. Thus, we might consider the average location of FaceFont-biased voxels to reflect the center of a face area within the left FG (l-FFA) and the average locus of the KoreanFont-biased voxels to reflect the center of the VWFA. If both FaceFont and KoreanFont training increase activation in “VWFA” voxels, then it would increase the probability that a “VWFA” voxel (instead of an “FFA” voxel) will be identified in a contrast involving the trained font. In this case, the mean location of voxels identified by the FaceFont localizer in FF participants would shift in the lateroposterior direction, as would the mean location of voxels identified by the KoreanFont localizer in KF-trained participants.

DISCUSSION

Years of interdisciplinary research have significantly advanced our understanding of optimal methods for reading instruction (Beck, 2006; Vaughn & Linan-Thompson, 2004), orthographic and phonological variables that interact with word identification (Ziegler & Goswami, 2006; Frost, 1998; Henderson, 1982), and the neural substrates that support reading (Turkeltaub, Eden, Jones, & Zeffiro, 2002; Pugh et al., 1996). Reading educators and researchers have begun to reach consensus on a number of key principles; for example, the idea that phonological access is a universal principle of skilled reading (Ziegler & Goswami, 2006; Frost, 1998). However, core theoretical issues remain (McCandliss et al., 2003; Price & Devlin, 2003; Plaut, McClelland, Seidenberg, & Patterson, 1996; Coltheart, Curtis, Atkins, & Haller, 1993), and for individuals with acquired and developmental reading disorders, there is a need for a better understanding of how the brain accommodates to the challenges associated with learning a writing system that connects with language functions and the factors that may modulate success (Lindgren, de Renzi, & Richman, 1985).

In literate adults, left-lateralized responses to orthographic stimuli are typically observed within the FG, at a location that has been termed the VWFA (Ben-Shachar et al., 2007; Polk et al., 2002; Cohen et al., 2000). Acquired and developmental reading disorders have been associated with dysfunction in this area as well (Shaywitz & Shaywitz, 2005; Cohen et al., 2003). While numerous findings point toward the critical role of the left FG in orthographic processing, there are several points of debate. These include which portions of the left FG are unimodal (Price & Devlin, 2003), uniquely activated by orthographic stimuli (Ben-Shachar et al., 2007; Price & Devlin, 2003), and specialized for different aspects of orthographic-to-phonological transformation (Vinckier et al., 2007; Pugh et al., 2000). Here we addressed the issue of why orthographic processing typically elicits a lateralized response within the left FG. We considered the linguistic bridge and visual perceptual accounts for left lateralization.

Evidence in Favor of a Linguistic Bridge Account

In the experiments described here, we identified lateralized training effects in a left FG region, in the vicinity of the expected location of the VWFA. The fact that training effects were found within this region for both KoreanFont and FaceFont favors the linguistic bridge account for the lateralization of orthographic processing. Graphemes with little resemblance to letters (i.e., FaceFont graphemes) engaged the left FG during reading, even when the perceptual characteristics of the graphemes would typically invoke a lateralized response in the right FG. Furthermore, this linguistic engagement of the left FG appears to be behaviorally relevant based on our finding that faster reading times (shorter latencies) correlated with greater activation in the left FG region for the participant-specific training font. Interpreting our findings in favor of a linguistic bridge account is consistent with recent work in which the left posterior occipitotemporal sulcus response to visual objects was modulated by tasks designed to encourage or suppress phonological encoding (Mano et al., 2013).

As an alternative explanation for our findings, it is possible that participants adopted an orthographic learning approach during training. For example, using a “remapping” strategy, participants could have associated each artificial grapheme with a specific English letter (e.g., the FaceFont grapheme for /k/ could be mapped onto the letter “k”). Such a strategy might allow FaceFont (or KoreanFont) to be read via indirect access to the neural substrates that support English orthographic processing. Circumstantial evidence argues against this interpretation of our findings. We made every attempt to avoid a link between artificial graphemes and English letters. When we listened to participants reading FaceFont and KoreanFont words and stories aloud, we did not hear evidence of a letter-naming strategy; instead, we heard many instances in which a participant decoded and blended the phonemes in a word based on the corresponding artificial graphemes. We also examined our word test naming data for evidence of a remapping strategy by splitting the items according to whether there was a mismatch between the number of phonemes in the word (i.e., the number of FaceFont or KoreanFont graphs needed to represent the word) and the number of letters in the printed English form of the word. There were no significant differences in the naming times for the two sets of items, in either the FF or the KF group (see Appendix C for a summary of results).

Another possibility is that the learned response to FaceFont (or KoreanFont) reflects some form of incidental processing that is irrelevant for reading the learned orthography. For instance, training may induce top-down attentional modulation of neurons within the left FG that had the preexisting capacity to respond to faces or prototypical orthographic forms. For FaceFont, it might be expected that similar attentional modulations would occur within the right FG at or near the FFA. We found no evidence for a training effect in the right FG. Furthermore, an incidental attention account cannot easily explain the shifts in the location of the learned versus unlearned responses to FaceFont and KoreanFont nor explain the observed relationship between the magnitude of the learned response and our behavioral measure of reading skill.

Additional evidence that the left FG is essential for the acquisition and use of FaceFont comes from a converging study involving an individual (AA1) with acquired alexia resulting from damage to the left FG (Moore, Brendel, & Fiez, in press). Across multiple attempts, AA1 failed to learn more than five face-phoneme mappings. Furthermore, she was unable to reliably decode English words that were printed in FaceFont (comprising only graphemes that she learned successfully). AA1 retained the ability to accurately name printed English letters and used her preserved letter recognition abilities to implement a letter-by-letter reading strategy that is the hallmark of acquired alexia. If FaceFont reading also rests on such a strategy, it would be expected that AA1 could use her preserved letter recognition abilities to learn FaceFont.

Future work will be needed to test the limits of the linguistic bridge account. Interestingly, although AA1 was unable to acquire the FaceFont orthography, she was able to learn a protosyllabary comprising 15 face-syllable mappings and then use the learned mappings to decode English words. These results indicate that nonalphabetic orthographies may reduce reliance on the left FG, potentially because they impose different phonological demands or recruit different types of visuospatial analysis. Consistent with this interpretation, it is notable that in readers of Chinese (a morphosyllabic writing system), visual presentation of Chinese characters produces bilateral activation in the FG in the vicinity of the VWFA and its potential homologue in the right hemisphere (Bolger, Perfetti, & Schneider, 2005). This unusual pattern of orthographic processing is found in both native readers of Chinese and in English-Chinese bilinguals (Nelson, Liu, Fiez, & Perfetti, 2009).

Perceptual Influences on Orthographic Processing

Although our results support the linguistic bridge account for the typical left lateralization of orthographic processing, there are indicators that the perceptual characteristics of the training fonts exerted an important influence. FF participants had slower reading latencies and less activation within the left FG compared with KF participants, pointing to an “orthographic bottleneck” for FF participants compared with KF participants, that is, a perceptual rather than linguistic cause for the poorer reading fluency attained by FF participants during training. We found subtle differences in the mean location of the peak responses to FaceFont and KoreanFont, which were further affected by training. These differences lead to differing hypotheses about the underlying neuronal causes for this

orthographic bottleneck. Each hypothesis is consistent with different aspects of our findings, so further research will be needed to distinguish between them.

One hypothesis is that acquisition of FaceFont recruits a pool of face-responsive neurons that lie near the vicinity of the VWFA, but not within VWFA. This would explain why the mean location of the learned response to FaceFont differs from the mean location of the learned response to KoreanFont. It is also consistent with findings from other studies, which indicate that differentiated or semidifferentiated neuronal pools exist in the vicinity of the VWFA (and its potential homologue in the right hemisphere; Downing, Chan, Peelen, Dodds, & Kanwisher, 2006). For instance, one recent study localized the responses to both faces and English words within both the left and right FG of individual participants. As expected, responses to words in the left FG were robust and located near the prototypical location of the VWFA. Weaker responses to faces could be found in the left FG in a subset of participants (2/4). Within an individual, the peak response for faces tended to be more anteromedial than the peak response for words. Across participants, however, there was substantial overlap between the face and word peak locations (Kawabata Duncan & Devlin, 2011). In the right FG, the same biases in the relative location of face versus word peak responses were observed, though as expected the responses to faces were more robust and consistent across participants than the responses to words. In this view, FaceFont may confront an orthographic bottleneck because it is forced to rely on a pool of face-sensitive neurons within the left FG that is less equipped for orthographic processing. The available pool may be small, and this may impose resource constraints; the pool may have a pattern of connections that make it a less suitable linguistic bridge, or the recruitment of face-responsive neurons may induce competition from typical face recognition processes.

An alternative hypothesis is that the acquisition of FaceFont necessarily recruits neurons that lie within the VWFA. Only this portion of the left FG, for instance, may have the innate or learned connectional architecture that is needed to form a suitable linguistic bridge. This hypothesis can explain why the mean location of the learned response to FaceFont does not differ from the unlearned response to KoreanFont. It is also consistent with evidence that neurons within a functionally localized VWFA have the capacity to respond to faces (Nestor, Behrmann, & Plaut, 2012; Dehaene et al., 2010). To explain this fact, some have suggested that the VWFA comprises interdigitated pools of neurons with different innate or learned response biases (e.g., neurons differentially tuned to respond to orthographic forms vs. objects or faces; Mano et al., 2013; Nestor et al., 2012). In this view, the boundaries of the VWFA delineate the tissue where there is a preponderance of orthographically biased neurons. Others have argued that the neurons within the VWFA have the innate capacity to respond to many types of visual stimuli, but the acquisition of literacy drives them to become tuned for the representation of orthographic forms (Dehaene et al., 2010). In this view, the boundaries of the VWFA delineate the tissue that was substantially influenced by this experience-dependent perceptual differentiation.

Irrespective of their differences, both of these views of the organization of the VWFA predict that there should be noisy response preferences within VWFA voxels and a loose form of topographic organization, which as noted earlier can explain the posterior-lateral shifts that were observed for the learned responses to both FaceFont and KoreanFont.

Importantly, either view can also be used to explain an orthographic bottleneck for FaceFont as compared with KoreanFont. In one case, the overall pool of face-biased neurons within the VWFA would be smaller than the overall pool of orthographically biased neurons, and this could impose resource constraints on the orthographic processing of FaceFont. In the other case, FaceFont might be forced to rely on less perceptually differentiated neurons at the margins of the VWFA or weakly responsive neurons with a preference for more prototypical orthographic stimuli.

Conclusions

In conclusion, this study demonstrates that neural tissue within the vicinity of the VWFA has the flexibility to respond to a range of “letters” or units in a writing system. The left-lateralized response to FaceFont in trained participants supports a linguistic bridge account for the lateralization of orthographic processing. For an alphabetic orthography, lateralized recruitment of neural tissue within the left FG appears to be governed by the demands of connecting visual processing with components of the speech system to permit the construction of a phonological representation that allows access to stored word knowledge. At the same time, the visual perceptual characteristics of a writing system do contribute to the specific location of activation within the left FG as well as to reading performance. Orthographic systems with atypical graphemes (like faces) appear to be susceptible to predispositions that result in a bottleneck in performance. These findings reinforce prior work that has implicated the left FG, the VWFA in particular, as a critical neuronal territory for the acquisition of skilled reading, and they suggest that reduced neuronal capacity in this territory could create an orthographic processing bottleneck that would slow word identification and access to stored word knowledge.

Acknowledgments

This project was funded by National Institutes of Health grant 1R01HD060388. A sincere thank you to Gal Ben-Yehuda, Jenna El-Wagaa, Alison Masey, Emma Anthony, Marge Gibson, Deborah Viszlay, Scott Kurdilla, and other members of the Fiez Lab for general assistance and helpful discussions related to this project.

References

- Arguin M, Bub D. Pure alexia: Attempted rehabilitation and its implications for the interpretation of the deficit. *Brain and Language*. 1994; 47:233–268. [PubMed: 7953616]
- Beck, IL. Making sense of phonics: The hows and whys. New York: Guilford Press; 2006.
- Behrmann M, Black SE, Bub D. The evolution of pure alexia: A longitudinal study of recovery. *Brain and Language*. 1990; 39:405–427. [PubMed: 2285860]
- Behrmann M, McLeod J. Rehabilitation for pure alexia: Efficacy of therapy and implications for models of normal word recognition. *Neuropsychological Rehabilitation*. 1995; 5:149–180.
- Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA. Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cerebral Cortex*. 2007; 17:1604–1641. [PubMed: 16956978]
- Black, SE.; Behrmann, M. Localization in alexia. In: Kertesz, A., editor. *Localization and neuroimaging in neuropsychology*. San Diego, CA: Academic Press; 1994. p. 331-376.
- Bolger DJ, Perfetti CA, Schneider W. Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping*. 2005; 25:92–104. [PubMed: 15846818]
- Changizi MA, Shimojo S. Character complexity and redundancy in writing systems over human history. *Proceedings of the Biological Sciences*. 2005; 272:267–275.

- Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lamberz G, Henaff MA, et al. The visual word form area. Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*. 2000; 123:291–307. [PubMed: 10648437]
- Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S. Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*. 2002; 125:1054–1069. [PubMed: 11960895]
- Cohen L, Martinaud O, Lemer C, Lehericy S, Samson Y, Obadia M, et al. Visual word recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*. 2003; 13:1313–1333. [PubMed: 14615297]
- Coltheart M, Curtis B, Atkins P, Haller M. Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*. 1993; 100:589–608.
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*. 1996; 29:162–173. [PubMed: 8812068]
- Daniel MS, Bolter JF, Long CJ. Remediation of alexia without agraphia: A case study. *Brain Injury*. 1992; 6:529–542. [PubMed: 1393187]
- Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, et al. How learning to read changes the cortical networks for vision and language. *Science*. 2010; 330:1359–1364. [PubMed: 21071632]
- Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N. Domain specificity in visual cortex. *Cerebral Cortex*. 2006; 16:1453–1461. [PubMed: 16339084]
- Ellis AW, Ansorge L, Lavidor M. Words, hemispheres, and processing mechanisms: A response to Marsolek and Deason. *Brain and Language*. 2007; 103:308–312.
- Fissell K, Tseytlin E, Cunningham D, Iyer K, Carter CS, Schneider W, et al. Fiswidgets: A graphical computing environment for neuroimaging analysis. *Neuroinformatics*. 2003; 1:111–125. [PubMed: 15055396]
- Fox, PT.; Pardo, JV. Does inter-subject variability in cortical functional organization increase with neural “distance” from the periphery?. In: Chadwick, DJ.; Whelan, J., editors. *Exploring brain functional anatomy with positron emission tomography*; Ciba Foundation Symposium; New York: Wiley; 1991. p. 125-144.
- Frost R. Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*. 1998; 123:71–99. [PubMed: 9461854]
- Gaillard R, Naccache L, Pinel P, Clemenceau S, Volle E, Hasboun D, et al. Direct intracranial, fMRI, and lesion evidence for the causal role of left interotemporal cortex in reading. *Neuron*. 2006; 50:191–204. [PubMed: 16630832]
- Gaydos, N. *Now I’m reading: For beginning readers*. Norwalk, CT: Innovative Kids; 2003.
- Goffaux V, Rossion B. Faces are “spatial”-holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception & Performance*. 2006; 32:1023–1039. [PubMed: 16846295]
- Grill-Spector K, Knouf N, Kanwisher N. The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*. 2004; 7:555–562.
- Hasson U, Levy I, Behrmann M, Hendler T, Malach R. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*. 2002; 34:479–490. [PubMed: 11988177]
- Henderson, L. *Orthography and word recognition in reading*. New York: Academic Press; 1982.
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*. 1997; 17:4302–4311. [PubMed: 9151747]
- Kanwisher N, Yovel G. The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions: Biological Sciences*. 2006; 361:2109–2128. [PubMed: 17118927]
- Kawabata Duncan KJ, Devlin JT. Improving the reliability of functional localizers. *Neuroimage*. 2011; 57:1022–1030. [PubMed: 21600292]
- Lindgren SD, de Renzi E, Richman LC. Cross-national comparisons of developmental dyslexia in Italy and the United States. *Child Development*. 1985; 56:1401–1417.

- Maher LM, Clayton MC, Barrett AM, Schober-Peterson D, Gonzalez-Rothi LJ. Rehabilitation of a case of pure alexia: Exploiting residual abilities. *Journal of the International Neuropsychological Society*. 1998; 4:636–647. [PubMed: 10050368]
- Mano QR, Humphries C, Desai RH, Seidenberg MS, Osmon DC, Stengel BC, et al. The role of left occipitotemporal cortex in reading: Reconciling stimulus, task, and lexicality effects. *Cerebral Cortex*. 2013; 23:988–1001. [PubMed: 22505661]
- Marsolek CJ, Deason RG. Hemispheric asymmetries in visual word-form processing: Progress, conflict, and evaluating theories. *Brain and Language*. 2007; 103:304–307. [PubMed: 17418887]
- McCandliss BD, Cohen L, Dehaene S. The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Science*. 2003; 7:293–299.
- Mecacci L. On spatial frequencies and cerebral hemispheres: Some remarks from the electrophysiological and neuropsychological points of view. *Brain and Cognition*. 1993; 22:199–212. [PubMed: 8373573]
- Melamed F, Zaidel E. Language and task effects on lateralized word recognition. *Brain and Language*. 1993; 45:70–85. [PubMed: 8353731]
- Montant M, Behrmann M. Pure alexia. *Neurocase*. 2000; 6:265–294.
- Moore MW, Brendel CB, Fiez JA. Reading faces: Investigating the use of a novel face-based orthography in acquired alexia. *Brain and Language*. (in press).
- Nelson JR, Liu Y, Fiez J, Perfetti CA. Assimilation and accommodation patterns in ventral occipitotemporal cortex in learning a second writing system. *Human Brain Mapping*. 2009; 30:810–820. [PubMed: 18381767]
- Nestor A, Behrmann M, Plaut DC. The neural basis of visual word form processing: A multivariate investigation. *Cerebral Cortex*. 2012; 23:1673–1684. [PubMed: 22693338]
- Perfetti, CA.; Hart, L. The lexical quality hypothesis. In: Verhoeven, L.; Elbro, C.; Reitsma, P., editors. *Studies in written language and literacy: Precursors of functional literacy*. Amsterdam/Philadelphia: John Benjamins Publishing Company; 2002. p. 189-214.
- Plaut DC, McClelland JF, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*. 1996; 103:56–115. [PubMed: 8650300]
- Polk TA, Stallcup M, Aguirre GK, Alsup DC. Neural specialization for letter recognition. *Journal of Cognitive Neuroscience*. 2002; 14:145–159. [PubMed: 11970782]
- Price CJ, Devlin JT. The myth of the visual word form area. *Neuroimage*. 2003; 19:473–481. [PubMed: 12880781]
- Pugh KR, Mencl WE, Jenner AR, Katz L, Frost SJ, Lee JR, et al. Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Review*. 2000; 6:207–213.
- Pugh KR, Shaywitz BA, Shaywitz SE, Constable RT, Skudlarski P, Fulbright RK, et al. Cerebral organization of component processes in reading. *Brain*. 1996; 119:1221–1238. [PubMed: 8813285]
- Reinke K, Fernandes M, Schwindt G, O'Craven K, Grady CL. Functional specificity of the visual word form area: General activation for words and symbols but specific network activation for words. *Brain and Language*. 2008; 104:180–189. [PubMed: 17531309]
- Schmahmann, JD.; Doyon, J.; Toga, AW.; Petrides, M.; Evans, AC. *MRI atlas of the human cerebellum*. San Diego, CA: Academic Press; 2000.
- Schneider, W.; Eschman, A.; Zuccolotto, A. *E-prime reference guide*. Pittsburgh: Psychology Software Tools Inc; 2002.
- Seki A, Koeda T, Sugihara S, Kamba M, Hirata Y, Ogawa T, et al. A functional magnetic resonance imaging study during sentence reading in Japanese dyslexic children. *Brain & Development*. 2001; 23:312–316. [PubMed: 11504602]
- Shaywitz SE, Shaywitz BA. Dyslexia (specific reading disability). *Biological Psychiatry*. 2005; 57:1301–1309. [PubMed: 15950002]
- Shtyrov Y, Pihko E, Pulvermuller F. Determinants of dominance: Is language laterality explained by physical or linguistic features of speech? *Neuroimage*. 2005; 27:37–47. [PubMed: 16023039]

Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain: An approach to medical cerebral imaging. Stuttgart: Thieme; 1988.

Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, et al. The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*. 2009; 168:242–249. [PubMed: 19564050]

Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA. Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage*. 2002; 16:765–780. [PubMed: 12169260]

Vaughn, S.; Linan-Thompson, S. Research-based methods of reading instruction, grades K-3. Alexandria, VA: Association for Supervision and Curriculum Development; 2004.

Vigneau M, Jobard G, Mazoyer B, Tzourio-Mazoyer N. Word and non-word reading: What role for the visual word form area? *Neuroimage*. 2005; 27:694–705. [PubMed: 15961322]

Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L. Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual-word form system. *Neuron*. 2007; 55:143–156. [PubMed: 17610823]

Wiederholt, JL.; Bryant, BR. The Gray Oral Reading Test, Fourth Edition (GORT-4). Austin, TX: Pro-Ed; 2001.

Woodhead ZV, Wise RJ, Sereno M, Leech R. Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. *Cerebral Cortex*. 2011; 21:2307–2312. [PubMed: 21368088]

Woods RP, Cherry SR, Mazziotta JC. Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*. 1992; 16:620–633. [PubMed: 1629424]

Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*. 1993; 17:536–546. [PubMed: 8331222]

Ziegler JC, Goswami U. Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*. 2006; 9:429–436. [PubMed: 16911438]

APPENDIX A

The FaceFont and KoreanFont corpora. The training font corpora are listed as well as the corresponding English phonemes (represented using the International Phonetic Alphabet) and examples of the sounds in English words. *Note:* Due to publication restrictions, only 5 of the 35 faces are permitted for use in print.

FaceFont								
KoreanFont	ㄱ	ㅋ	ㆁ	ㅁ	ㄱ	ㅋ	ㅁ	ㅁ
Phoneme(s)	/a, ɔ/	/e/	/aɪ/	/b/	/ɪʃ/	/d/	/eɪ/	/f/
Example(s)	hat ball	cat had	by pie	bud	watch chair	den	bait base	bet head fix
FaceFont								
KoreanFont	ㄱ	ㅋ	ㄴ	ㄷ	ㅌ	ㅍ	ㅑ	ㅓ
Phoneme(s)	/g/	/h/	/ɪ/	/ɪ/	/dʒ/	/j/	/k/	/l/
Example(s)	go	hit	he eat	hit him	judge	you yarn	cake	lamp my
FaceFont								
KoreanFont	ㅍ	ㄱ	ㄱ	ㅇ	ㅌ	ㅋ	ㅍ	ㅑ
Phoneme(s)	/n/	/ŋ/	/o/	/oʊ/	/u/	/aʊ/	/p/	/r, ɹ/
Example(s)	no	ring	no boat	toy	look could	cow house	got	face work soap
FaceFont								
KoreanFont	ㅍ	ㅍ	ㅍ	ㅇ	ㅍ	ㅍ	ㅍ	ㅍ
Phoneme(s)	/ʃ, ʒ/	/t/	/θ, ð/	/u/	/s, ɔ/	/v/	/w/	/z/
Example(s)	shirt measure	top	thumb this	tune food	cut the	van	wash	zip music

APPENDIX B

PubMed Search Results for Publications in 2012 that Reported Specific Coordinate Loci Associated with the VWFA

2012 Publication	TCX	TCY	TCZ	Analysis Type	Contrast
Striem-Amit et al., <i>Neuron</i>	-45	-58	-5	Group	Letters vs. Faces/Houses/Body Shapes/ Objects/Textures
Song et al., <i>Journal of Neuroscience</i>	-42	-51	-15	Individual	Chinese characters vs. landscapes
Monzalvao et al., <i>Neuroimage</i>	-45	-45	-15	Group	Normal readers > Dyslexics for Word vs. Rest contrast
Park et al., <i>PLoS One</i>	-40	-43	-14	Group	Words/Pseudowords/Consonant Strings vs. Numbers
Xu et al., <i>Neuropsychologia</i>	-44	-56	-9	Individual	Words vs. Faces/Textures
Mean	-43	-51	-12		
SD	±2	±7	±4		

APPENDIX C

To look for evidence of a remapping strategy, we assessed naming latencies for real word items that contained three or four artificial graphs in Word Test 9 (only correct responses were included). Words with 0, 1, or 2 letter–phoneme mismatches were included in the analysis. For example, the word “fight” comprises five printed English letters and three phonemes (therefore, three artificial graphs), for a letter–phoneme mismatch of 2. In a 2×3 between-subject ANOVA with Training Group (FF-trained, KF-trained) and Letter–Phoneme Mismatch (0, 1, 2) as factors, there was a main effect of Training Group, as expected, $F(1, 620) = 92.14, p < .001$. However, there was no main effect of Letter–Phoneme Mismatch ($p = .23$) and no Group \times Letter–Phoneme Mismatch interaction ($p = .25$).

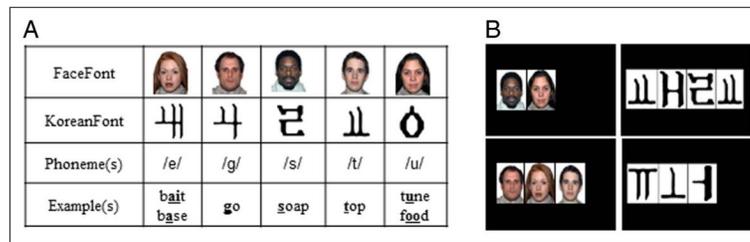


Figure 1. Examples of FaceFont and KoreanFont. (A) A sample of five FaceFont and KoreanFont “letters,” their corresponding English phonemes (represented using the International Phonetic Alphabet), and examples of the sounds in English words (see Appendix A for full corpora). (B) Examples of word level presentation from upper left moving clockwise: “sue,” “test,” “jar,” “gate.” Note: The FaceFont words shown in (B) were not used in the actual training or testing. Due to publication restrictions, only 5 of the 35 faces are permitted for use in print; the words shown here were derived from these five faces for example only.

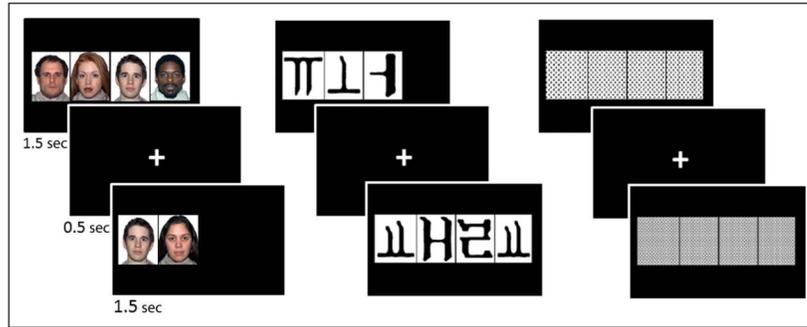


Figure 2.
Examples of FaceFont words, KoreanFont words, and pattern stimuli presented during the fMRI session.

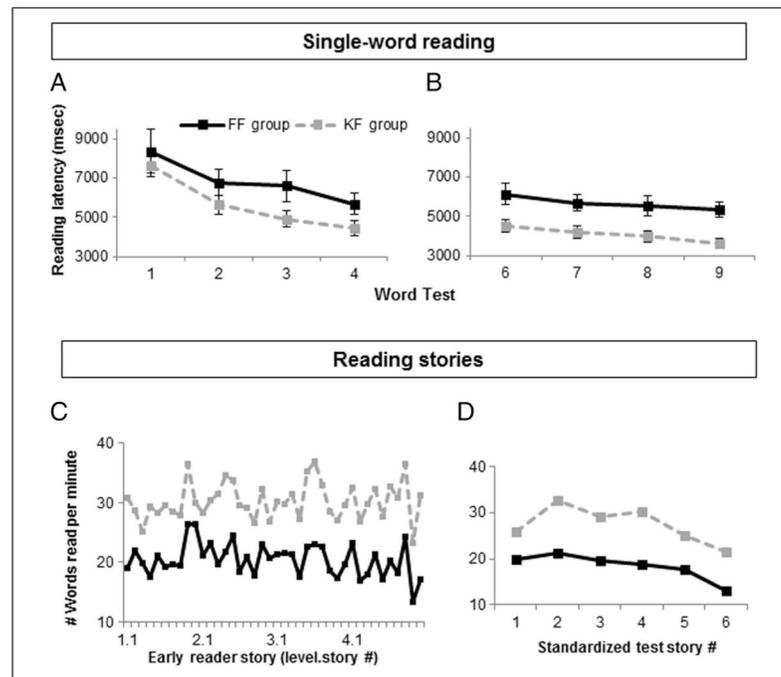


Figure 3. Reading latencies. Reading latency (msec) on single-word reading tests after (A) word level training in Week 1 and (B) story level training in Week 2. Latencies are reported for correct responses only and include words and nonwords. Both training groups improved in reading speed. FF participants improved at a similar rate as KF participants, although their average latency times were slower. The number of WPM was computed for reading performance on (C) the early reader training stories and (D) the standardized reading test stories. Graphs are consistent with single-word reading performance—the pattern of performance is similar between the two groups, although the FF group is slower.

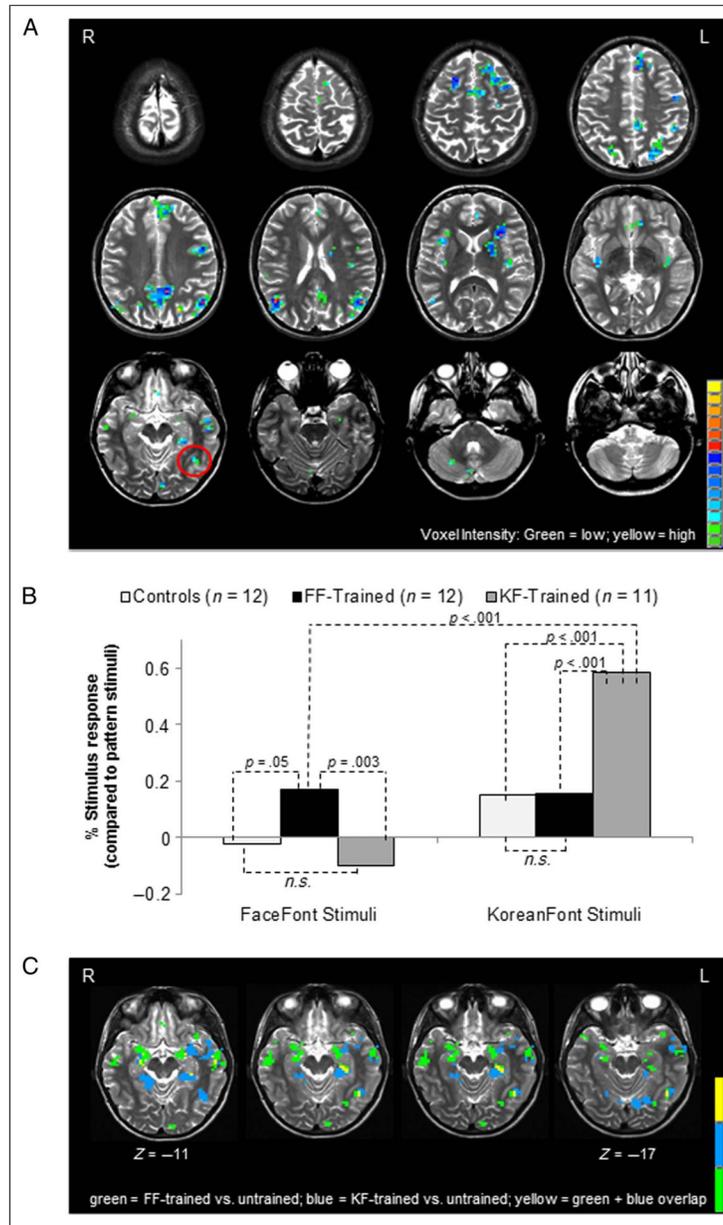


Figure 4. Neural response to training. A 2×2 voxel-wise ANOVA (Training Group \times Font) was computed (untrained controls and pattern stimuli were not included in the ANOVA). With a contiguity threshold = 5 and a corrected $p = .01$, a significant Training Group \times Font interaction was found in the left FG that is consistent with the VWFA (peak voxel: $-40, -50, -11$). (A) Whole brain map showing the activation pattern of the Training Group \times Font interaction. Activation in the left FG is circled in red. (B) Simple effects analyses showed that the trained groups had a significantly greater response in the LFG-VC when viewing their trained font compared with when the untrained groups viewed their trained font. When viewing the FaceFont stimuli: FF group response vs. KF group response, $p = .003$; FF group response vs. control response, $p = .05$; no significant differences between the

two untrained groups, KF group vs. controls, $p > .05$. When viewing KoreanFont stimuli: KF group response vs. FF group response, $p < .001$; KF group response vs. control response, $p < .001$; no significant differences between the two untrained groups, FF group response vs. controls, $p > .05$. The magnitude of the trained font response was significantly greater for the KF group compared with the FF group, $p < .001$. (C) Map of the left FG showing the activation pattern from a secondary conjunction analysis comparing the response to FaceFont in trained versus untrained (KF group and controls) participants with the response to KoreanFont in trained versus untrained (FF group and controls) participants. Results converge with the results in (A) and (B). Training effects for both fonts were found within the LFG, and the identified voxel clusters partially overlapped at a locus at or near the typical location of the VWFA. There was an absence of training effects in the right FG.

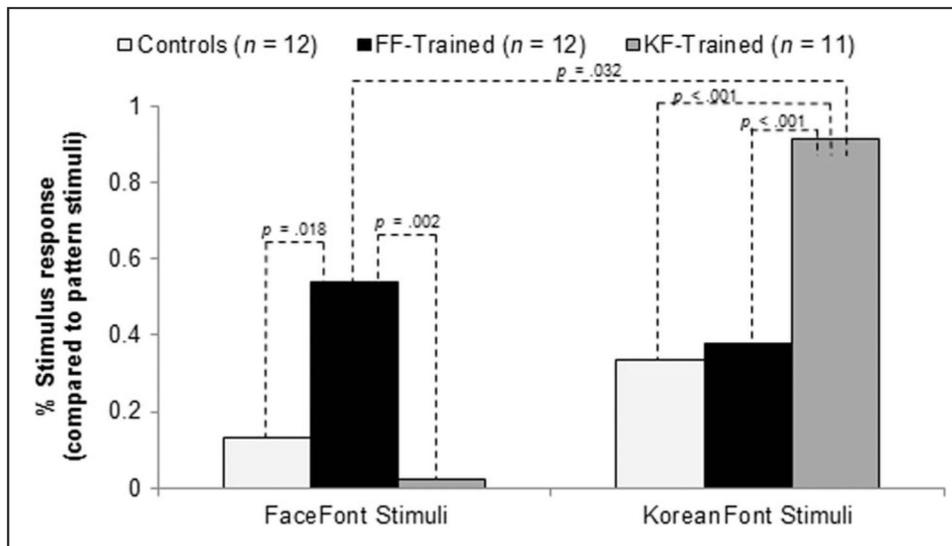


Figure 5.

Activation of KF-localized voxels. A 2×2 mixed ANOVA (Training Group \times Font) was computed (untrained controls and pattern stimuli were not included in the ANOVA). There was a Training Group \times Font interaction ($p < .001$) and a main effect of Font ($p < .001$), but no significant main effect of Training Group ($p = .932$). Significant independent sample t tests are shown. The trained groups showed a significantly greater response in the KF-localized voxels when viewing their trained font compared with when the untrained groups viewed the font. When viewing the FaceFont stimuli: FF group response vs. KF group response, $p = .002$; FF group response vs. control response, $p = .018$; no significant differences between the two untrained groups, KF group vs. controls, $p = .373$. When viewing KoreanFont stimuli: KF group response vs. FF group response, $p < .001$; KF group response vs. control response, $p < .001$; no significant differences between the two untrained groups, FF group response vs. controls, $p = .628$. The magnitude of the trained font response was significantly greater for the KF group compared with the FF group, $p = .032$.

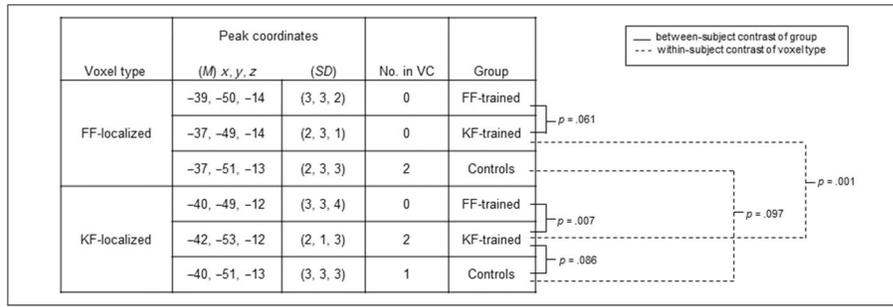


Figure 6.

A 2×2 mixed MANOVA (Training Group \times Voxel Type) was computed to compare differences in the average location of peak responses (control participants were not included in the mixed MANOVA). There was a main effect of Voxel Type ($p < .001$) and a Group \times Voxel Type interaction ($p = .001$), but no main effect of Group ($p = .339$). Simple MANOVA contrasts are represented in the figure. There were no significant differences between the groups' peak voxel coordinates for the FF-localized voxels: FF-trained vs. KF-trained, $p = .061$; FF-trained vs. control, $p = .375$; KF-trained vs. control, $p = .414$. For the KF-localized voxels, there was a significant difference in peak voxel coordinates between the trained groups: FF-trained vs. KF-trained peak voxel, $p = .007$. There were no other significant simple contrast differences using the KF-localized voxels: FF-trained vs. control peak voxel, $p = .460$; KF-trained vs. control peak voxel, $p = .086$. When comparing the FF- and KF-localized coordinates within a group, significant differences were found in the KF-trained group ($p = .001$); there were no significant differences between voxel type found for the FF-trained group ($p = .212$) or the control group ($p = .097$). No. in VC = number of identified individual peaks that fall within the interaction voxel cluster from the primary analysis. The limited number of individual peaks within the VC is not surprising because the VC was smaller than the typical extent of variability seen in individual response peak (Fox & Pardo, 1991).

Table 1

Daily Schedule for Trained Participants

<i>Week</i>	<i>Session</i>	<i>Tasks</i>	
Week 1	Session 1	Phoneme Training Phoneme Test	
	Session 2	Phoneme Review Word Level Training Word Test (1)	
	Sessions 3–5	Word Level Training Word Test (2–4)	
	Week 2	Sessions 6	Word Test (5) Supplemental Word Test (5S) Story Level Training Word Test (6) Supplemental Word Test (6S)
		Sessions 7–9	Story Level Training Word Test (7–9) Supplemental Word Test (7S–9S)
Session 10		Reading Test (GORT-4) fMRI	

Table 2

Coordinates of Peak Activation within Clusters that Showed a Significant Main Effect of Font

<i>Direction of Effect</i>	<i>Cluster Location</i>	<i>Peak Coordinates</i>	<i>Cluster Size (No. of Voxels)</i>
FaceFont > KoreanFont	Right FG	23, -88, -14	32
	Right FG	34, -37, -21	6
	Left cerebellum, VI	-13, -91, -18	7
KoreanFont > FaceFont	Right supramarginal gyrus	48, -34, 36	6
	Left cuneus	-23, -75, 30	5
	Left FG	-43, -56, -7	9

A 2×2 voxel-wise ANOVA (Training Group \times Font) was computed with a contiguity threshold = 5 and a corrected $p = .01$ (uncorrected $p = .000005$). Direction of the main effect was based on a normal scale that included both positive and negative signal activations. In the one case of negative signal activations for both fonts (L Cerebellum, VI), the effect direction indicates that the FaceFont stimuli showed less suppression than the KoreanFont stimuli (i.e., FaceFont % stimulus response was less negative than KoreanFont).