Decoding and disrupting left midfusiform gyrus activity during word reading

Elizabeth A. Hirshorn[a,1,2], Yuanning Li[b,c,d,1], Michael J. Ward[b], R. Mark Richardson[b,c], Julie A. Fiez[a,c,e], and Avniel Singh Ghuman[b,c,d]

[a]Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260; [b]Department of Neurological Surgery, University of Pittsburgh, Pittsburgh, PA 15260; [c]Center for the Neural Basis of Cognition, Carnegie Mellon University and University of Pittsburgh, Pittsburgh, PA 15213; [d]Program in Neural Computation, Center for the Neural Basis of Cognition, Carnegie Mellon University and University of Pittsburgh, Pittsburgh, PA 15213; and [e]Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260

Edited by Brian A. Wandell, Stanford University, Stanford, CA, and approved May 17, 2016 (received for review March 11, 2016)

The nature of the visual representation for words has been fiercely debated for over 150 y. We used direct brain stimulation, pre- and postsurgical behavioral measures, and intracranial electroencephalography to provide support for, and elaborate upon, the visual word form hypothesis. This hypothesis states that activity in the left midfusiform gyrus (lmFG) reflects visually organized information about words and word parts. In patients with electrodes placed directly in their lmFG, we found that disrupting lmFG activity through stimulation, and later surgical resection in one of the patients, led to impaired perception of whole words and letters. Furthermore, using machine-learning methods to analyze the electrophysiological data from these electrodes, we found that information contained in early lmFG activity was consistent with an orthographic similarity space. Finally, the lmFG contributed to at least two distinguishable stages of word processing, an early stage that reflects gist-level visual representation sensitive to orthographic statistics, and a later stage that reflects more precise representation sufficient for the individuation of orthographic word forms. These results provide strong support for the visual word form hypothesis and demonstrate that across time the lmFG is involved in multiple stages of orthographic representation.

fusiform gyrus | word reading | temporal dynamics | intracranial EEG | electrical stimulation

A central debate in understanding how we read, documented at least as far back as Charcot, Dejerine, and Wernicke, has revolved around whether visual representations of words can be found in the brain. Specifically, Charcot and Dejerine posited the existence of a center for the visual memory of words (1), whereas Wernicke firmly rejected that notion, proposing that reading only necessitates representations of visual letters that feed forward into the language system (2). Similarly, the modern debate revolves around whether there is a visual word form system that becomes specialized for the representation of orthographic knowledge (e.g., the visual forms of letter combinations, morphemes, and whole words) (1, 3, 4). One side of the debate is characterized by the view that the brain possesses a visual word form area that is "a major, reproducible site of orthographic knowledge" (5), whereas the other side disavows any need for reading-specific visual specialization, arguing instead for neurons that are "general purpose analyzers of visual forms" (6).

The visual word form hypothesis has attracted great scrutiny because the historical novelty of reading makes it highly unlikely that evolution has created a brain system specialized for reading; this places the analysis of visual word forms in stark contrast to other processes that are thought to have specialized neural systems, such as social, verbal language, or emotional processes, which can be seen in our evolutionary ancestors. Thus, testing the word form hypothesis is critical not only for understanding the neural basis of reading, but also for understanding how the brain organizes information that must be learned through extensive experience and for which we have no evolutionary bias.

Advances in neuroimaging and lesion mapping have focused the modern debate surrounding the visual word form hypothesis on the left midfusiform gyrus (lmFG). This focus reflects widespread agreement that the lmFG region plays a critical role in reading. Supporting evidence includes demonstrations that literacy shapes the functional specialization of the lmFG in children and adults (7–10); the lmFG is affected by orthographic training in adults (11, 12); and damage to the lmFG impairs visual word identification in literate adults (13, 14). However, debate remains about whether the lmFG constitutes a visual word form area (3, 5, 15–18) or not (6, 19, 20); that is, does it support the representation of orthographic knowledge about graphemes, their combinatorial statistics, orthographic similarities between words, and word identity (21), or does it have receptive properties tuned for general purpose visual analysis, with lexical knowledge emerging from the spoken language network (6)?

To test the limits of the modern visual word form hypothesis, we present results from four neurosurgical patients (P1–P4) with electrodes implanted in their lmFG. We acquired pre- and postsurgery neuropsychological data in P1, performed direct cortical stimulation in P1 and P2, and recorded intracranial electroencephalography (iEEG) in all four participants to examine a number of indicators that have been proposed as tests for the visual word form hypothesis by both supporters and opponents of this hypothesis (5, 6). Pattern classification methods

Significance

A central issue in the neurobiology of reading is a debate regarding the visual representation of words, particularly in the left midfusiform gyrus (lmFG). Direct neural recordings, electrical brain stimulation, and pre-/postsurgical neuropsychological testing provided strong evidence that the lmFG supports an orthographically specific "visual word form" system that becomes specialized for the representation of orthographic knowledge. Machine learning elucidated the dynamic role lmFG plays with an early processing stage organized by orthographic similarity and a later stage supporting individuation of single words. The results suggest that there is a dynamic shift from gist-level to individuated orthographic representation in the lmFG in service of visual word recognition.

Author contributions: E.A.H., R.M.R., J.A.F., and A.S.G. designed research; E.A.H., Y.L., M.J.W., R.M.R., and A.S.G. performed research; E.A.H., Y.L., M.J.W., J.A.F., and A.S.G. analyzed data; and E.A.H., Y.L., J.A.F., and A.S.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]E.A.H. and Y.L. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: hirshorn@pitt.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1604126113/-/DCSupplemental.

from machine learning were then used to measure whether neural coding in this region is sufficient to represent different aspects of orthographic knowledge, including the identity of a printed word. We separately evaluated the time course of lmFG sensitivity to different aspects of orthographic information to assess both early processing, which should exclusively or predominantly capture bottom-up visual processing, and later processing, which likely captures feedback and recurrent interactions with higher-level visual and nonvisual regions. Consequently, we were able to assess the dynamic nature of orthographic representation within the lmFG and thereby provide a novel perspective on the nature of visual word representation in the brain.

Results

Verification of Orthographic Selectivity at lmFG Electrode Sites. To identify their seizure foci, four patients with medically intractable epilepsy underwent iEEG, which included insertion of multicontact electrodes into or on their ventral temporal cortex (VT) (Fig. 1). To assess the word sensitivity and specificity of lmFG, we used a Gaussian naïve Bayes classifier to decode the neural activity (single trial potentials) while participants viewed three different categories of visual stimuli: words, bodies, and phase-scrambled objects (30 images per category, each repeated once). In each patient in electrode contacts in lmFG, we observed a strong early sensitivity to words at 100–400 ms (Fig. 2 *A* and *B*), which was verified using a classifier model (Fig. 2*C*; averaged peak $d' = 1.26$, at 245 ms after stimulus onset, $P < 0.001$; see Figs. S1–S4 for each individual contact on the electrodes from each participant). The position of the lmFG electrode contacts in the anterior end of the posterior fusiform sulcus is consistent with the putative visual word form area described in the functional neuroimaging literature (22–24). Further, the timing of the category selective response is consistent with evoked potential findings obtained from scalp electrodes (25) and previous iEEG studies (23, 26–28), which have described orthographic-specific effects ∼200 ms after stimulus onset.

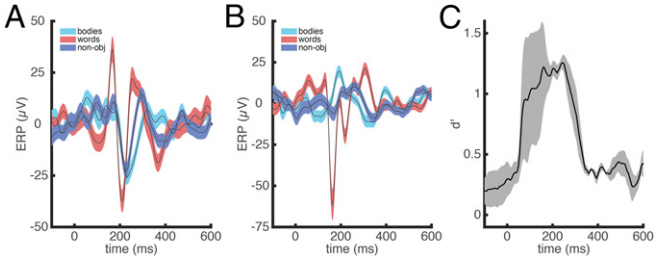

Fig. 2. Verification of orthographic selectivity at lmFG electrode site. (*A*) Example of averaged ERP across lmFG electrodes in one of the participants (P1) for three different stimulus categories (bodies, words, and nonobjects). The colored areas indicate SEs. (*B*) Averaged ERP across all lmFG electrodes and across all of the participants for three different stimulus categories (bodies, words, and nonobjects). The colored areas indicate SEs. (*C*) Time course of word categorical sensitivity in lmFG electrodes measured by sensitivity index d' (mean d' plotted against the beginning of the 100-ms sliding window), averaged across three participants. The MTPA classifier uses time-windowed single-trial potential signal from the electrodes from each subject (window length = 100 ms) with each time point in the window from each electrode as multivariate input features (see *Methods* for details). Across-participant SEs are shaded gray. See Figs. S1–S4 for single-electrode word categorical sensitivity.

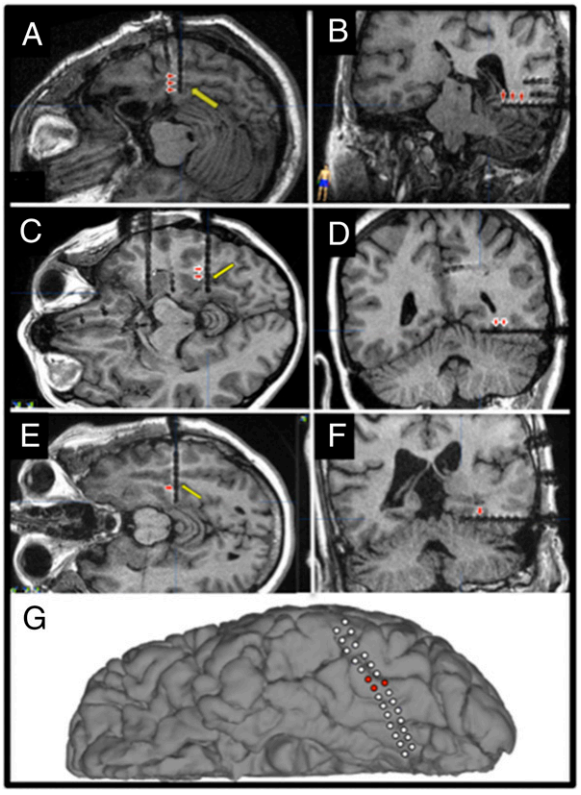

Fig. 1. Location of implanted electrode. Individual electrode contacts are visible on axial (*A*, *C*, and *E*) and coronal (*B*, *D*, and *F*) views and cortical reconstruction (*G*) of the postimplantation MRI (P1: *A* and *B*; P2: *C* and *D*; P3: *E* and *F*; P4: *G*). The VT depth electrodes were placed at the anterior end of the midfusiform sulcus in P1–P3 (yellow arrow), and P4 was implanted with a left temporal subdural grid crossing the lmFG. Red arrowheads (*A*–*F*) and red filled circles (*G*) indicate the word-selective contacts identified in the category localizer, which were used in subsequent electrophysiological and/or stimulation experiments. Talairach coordinates (*x*, *y*, *z*) corresponding to the word-selective contacts were located in postoperative MRI structural images, and were all identified in the left fusiform gyrus, BA 37 (P1 electrodes: −31, −36, −13; −35, −37, −13; −39, −38, −12; P2 electrodes: −30, −46, −11; −34, 6, −12; P3 electrodes: −31, −35, −14; P4 electrodes: −38, −51, −21; −41, −50, −22; −41, −54, −20).

After completion of the iEEG study, in P1 a focal resection in the posterior basal temporal lobe was performed, which included removal of tissue at the location of the implanted VT electrode (Fig. S5), leading us to predict that P1 would exhibit postsurgical changes in visual word recognition consistent with acquired alexia (13). Neuropsychological assessments of naming times were conducted pre- and postsurgery at 1.5 wk (acute), 6 wk, and 3 mo to assess the impact of the resection on his perception of visual stimuli. P1 was asked to name words (three, five, or seven letters) (14) and a mixed set of stimuli (words, letters, single digits, three-digit numbers, famous faces, objects, music notes, and guitar tabs) aloud as rapidly and accurately as possible. After removal of the area surrounding the VT electrode, P1 showed the characteristics of acquired alexia—specifically, letter-by-letter reading (Fig. 3*C*), and longer naming times, particularly for letters and words (Fig. 3*D*), as predicted based on the role of this area in orthographic processing (13, 14). Additionally, orthographic processes were impacted to a greater degree than phonological processes by the resection (Fig. S6). See *SI Results* for further description and elaboration on P1's postresection reading deficits.

The anatomical locus and category specificity of the recorded iEEG response in P1–P4, and the postresection alexia in P1, were highly consistent with our localization of lmFG electrodes to tissue that is central to the visual word form debate. We then tested specific putative indicators of the visual word form hypothesis using data obtained from cortical stimulation (P1 and P2) and iEEG (P1, P3, and P4) from these electrode sites.

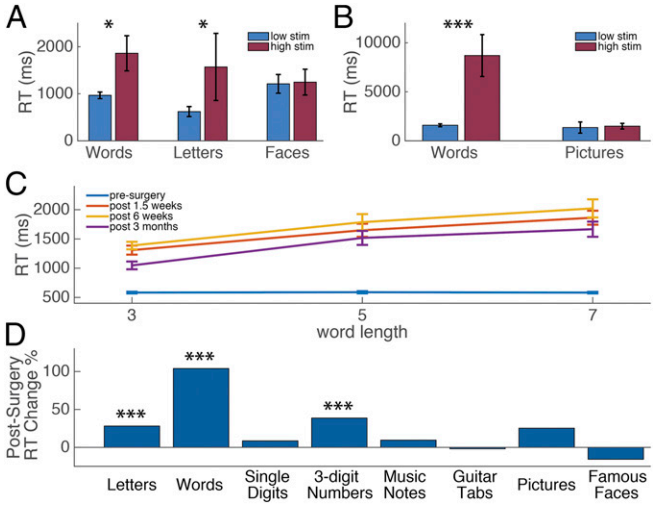

Fig. 3. The effect of stimulation on naming times in lmFG and pre- and postsurgery neuropsychological naming task performance. (*A*) The average naming reaction time for words, letters, and faces under low stimulation (1–5 mA) and high stimulation (6–10 mA) to lmFG electrodes in P1. Error bars correspond to SE, $*P < 0.05$. (*B*) The average naming reaction time for words and pictures under low stimulation (1–5 mA) and high stimulation (6–10 mA) to lmFG electrodes in P2. Error bars correspond to SE, $***P < 0.001$. (*C*) Word length effect pre- and postsurgery in P1. (*D*) Average percent change in reaction time in the mixed naming task pre- vs. postsurgery in P1, $***P < 0.001$.

Disrupting lmFG Activity Impairs Both Lexical and Sublexical Orthographic Processing. One indicator of whether the lmFG functions as a specialized visual word form system is whether disrupting its activity using electrical stimulation impairs the normal perception of both printed words and sublexical orthographic components (26, 27), but not other kinds of visual stimuli. As part of presurgical language mapping, P1 and P2 underwent an electrical stimulation session where they named two kinds of orthographic stimuli [words (P1 and P2) and letters (P1)], as well nonorthographic objects [faces (P1) and pictures (P2)]. We hypothesized that high stimulation (6–10 mA) to the lmFG electrodes would cause greater disruption to reading orthographic stimuli than low stimulation (1–5 mA) due to the observed category specificity of the iEEG response, but no disruption would be seen for stimulation during object (face or picture) naming. Indeed, P1 and P2 were significantly slower at reading words at high stimulation than low stimulation [Fig. 3 *A* and *B*; P1: mean $RT_{low\ stim}$ = 967 ms, mean $RT_{high\ stim}$ = 1,860 ms, $t(18)$ = 2.42, Cohen's d = 1.14, P = 0.026; P2: mean $RT_{low\ stim}$ = 1,586 ms, mean $RT_{high\ stim}$ = 8,700 ms, $t(7)$ = 11.28, Cohen's d = 5.15, P < 0.001]. P1 also misidentified 5% of words (naming "number" as "nature") under high stimulation on the lmFG electrodes. P2 did not misidentify any words, but was generally unable to name words until the stimulation had ceased. Her self-report suggested an orthographic disruption rather than speech arrest. Specifically, for the word "illegal," she reported thinking two different words at the same time, and trying to combine them. For the word "message," she reported thinking that there was an "N" in the word (Movie S1). P1 was also asked to name single letters during stimulation in lmFG electrodes. With limited letter trials during stimulation (two low stimulation and five high stimulation), there was no significant difference in reaction time in letter naming between high and low stimulation. However, P1 responded incorrectly to two letter stimuli, initially responding "A" for "X," and responding "F" and then "H" to the visual stimulus "C," both of which he had previously named accurately during the stimulation session (Movie S2). Importantly, naming times for nonorthographic stimuli were not significantly affected by stimulation in lmFG electrodes [P1, faces: mean $RT_{low\ stim}$ = 1,211 ms, mean $RT_{high\ stim}$ = 1,246 ms, $t(12)$ = 0.11, Cohen's d = 0.05, P = 0.92; P2, pictures: mean $RT_{low\ stim}$ = 1,350 ms, mean $RT_{high\ stim}$ = 1,490 ms, $t(10)$ = 0.18, Cohen's d = 0.13, P = 0.86]. (Naming times for pictures did not differ between low- and high-stimulation picture trials in P2 despite evidence of afterdischarges—abnormal activity that continues after stimulation is turned off—on three of four high-stimulation trials. No afterdischarges were seen during word naming.)

These results are consistent with previous reports of selective impairments due to stimulation in the lmFG for reading orthographic stimuli (29). Notably, the category-specific perceptual alteration seen in P1 and P2 reveals visual feature distortions that are similar to those reported for faces when stimulating right mFG (30). These stimulation results indicate that disruption of lmFG function impairs both the skilled identification of visual words and sublexical components of word forms (i.e., letters), supportive of the visual word form hypothesis.

Electrophysiological Evidence for a Visual Word Form Representation in the lmFG. We next used techniques from machine learning in iEEG data from P1 and P4 to assess the sensitivity of lmFG to sublexical, orthographic statistics (bigram frequency) that has been hypothesized as an indicator for a visual word form system (16, 21). To examine the dynamics of orthographic statistic sensitivity, we used a multivariate temporal pattern analysis (MTPA) classification procedure to test how the lmFG represents aspects of orthographic knowledge critical to the word form hypothesis at different stages of the time course.

To measure sublexical sensitivity as a test of the word form hypothesis, P1 and P4 performed a covert naming task with high- and low-bigram frequency words, controlled for lexical frequency. The MTPA classifier was sensitive to differences between high- and low-bigram frequency during a relatively early time window in both participants (Fig. 4; P1: peak accuracy = 58.6%, $P < 0.05$ at 200–330 ms after stimulus onset; P4: peak accuracy = 60.2%, $P < 0.05$ at 210–310 ms after stimulus onset; all classification analyses were tested using permutation tests to correct for multiple comparisons). This finding is consistent with early discrimination in the basal temporal cortex between words and pseudowords in Kanji, which differ in the likelihood and order of cooccurrence of two characters within a word (31). It has been noted that testing the visual word form hypothesis requires examining the representation in lmFG that results primarily from feedforward input from earlier parts of the ventral visual processing stream (5). Thus, the result that sublexical aspects of orthographic information begin at a

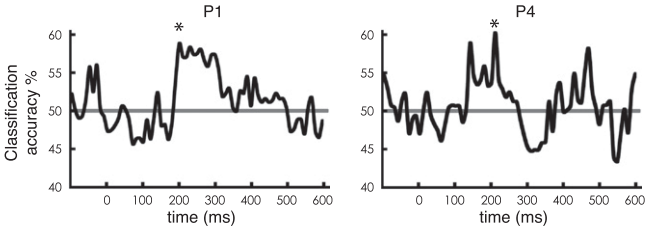

Fig. 4. Dynamics of sensitivity to sublexical orthographic statistics (bigram frequency) in the lmFG. Classification accuracy time course for comparison between low-bigram frequency real words (low BG) vs. high-bigram frequency real words (high BG) in lmFG electrodes for P1 and P4, respectively, plotted against the beginning of the 100-ms sliding window. The classifier uses time-windowed single-trial potential signal from the electrodes from each subject (window length = 100 ms) with each time point in the window from each electrode as multivariate input features (see *Methods* for details). The asterisk (*) corresponds to the peak of the windows in which $P < 0.05$ corrected for multiple comparisons. The P = 0.05 significance threshold corresponds to accuracy = 58.2% (P1) and 59.3% (P4). The horizontal gray line at 50% indicates chance level.

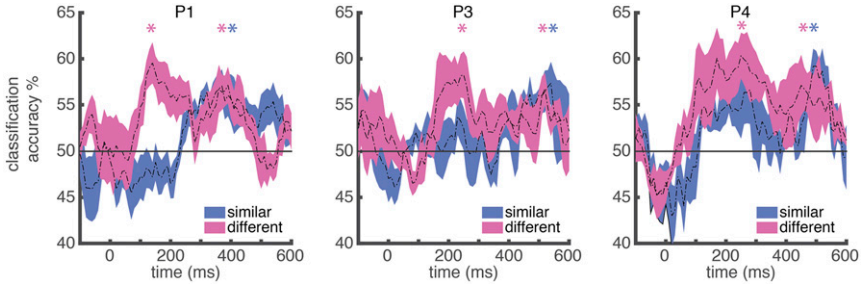

Fig. 5. Dynamics of word individuation selectivity in the lmFG. Dynamics of averaged pairwise word individuation accuracy for different conditions in lmFG electrodes for P1, P3, and P4, respectively, plotted against the beginning of the 100-ms sliding window. The classifier uses time-windowed single-trial potential signal from the electrodes from each subject (window length = 100 ms) with each time point in the window from each electrode as multivariate input features (see *Methods* for details). The time course of the accuracy is averaged across all word pairs of the corresponding conditions. The colored areas indicate SEs. Similar pair: a pair of words that have the same length and are only different in one letter, e.g., lint and hint. Different pair: a pair of words that have the same length and are different in all letters, e.g., lint and dome. Horizontal gray line indicates chance level (accuracy = 50%). Colored asterisk (*) corresponds to the peak of the windows in which $P < 0.05$ corrected for multiple comparisons. The $P = 0.05$ significance threshold corresponds to accuracy = 56.5% (P1), 56.0% (P3), and 57.1% (P4).

relatively early time point in processing is supportive of the word form hypothesis (5, 6, 16, 21, 32).

Temporal Dynamics of Word Individuation in lmFG. To further elucidate the dynamic nature of orthographic representation, we next looked at the sensitivity of lmFG to different aspects of individual words in P1, P3, and P4. Using words that varied in their degree of visual similarity (e.g., words that differed by one letter vs. all letters), we determined at what similarity level an MTPA classifier could discriminate between any two items. We found that at an early time window after stimulus onset, an MTPA classifier could significantly discriminate between words that did not share any letters (e.g., lint vs. dome; P1: peak classification accuracy = 59.6%, $P < 0.05$ from 120 to 250 ms; P3: peak classification accuracy = 58.3%, $P < 0.05$ from 180 to 360 ms; P4: peak classification accuracy = 60.3%, $P < 0.05$ from 100 to 430 ms, all P values were corrected for multiple time comparisons; Fig. 5), but could not discriminate between words that only differed by one letter (e.g., lint vs. hint; P1: peak classification accuracy = 52.7%, $P > 0.1$; P3: peak classification accuracy = 53.7%, $P > 0.1$; P4: peak classification accuracy = 56.6%, $P > 0.05$; Fig. 5). This result demonstrates an organization governed by an orthographic similarity space at the sublexical level, a finding consistent with our observation of bigram frequency effects in a relatively early time window. However, within a later time window, an MTPA classifier could discriminate between any two words (Fig. 5); notably, this includes word pairs with only one letter difference (P1: peak classification accuracy = 57.1%, $P < 0.05$ from 360 to 470 ms; P3: peak classification accuracy = 57.3%, $P < 0.05$ from 470 to 640 ms; P4: peak classification accuracy = 59.2%, $P < 0.05$ from 490 to 620 ms).

Discussion

Our findings, which indicate that orthographic representation within the lmFG qualitatively shifts over time, provide a novel advancement on the debate about the visual word form hypothesis (1, 2). Specifically, we demonstrated that lmFG meets all of the proposed criteria for a visual word form system: early activity in lmFG coded for orthographic information at the sublexical level, disrupting lmFG activity impaired both lexical and sublexical perception, and early activity reflected an orthographic similarity space (24). Early activity in lmFG is sufficient to support a gist-level representation of words that differentiates between words with different visual statistics (e.g., orthographic bigram frequency).

Notably, the results in the late time window suggest that orthographic representation in lmFG shifts from gist-level representations to more precise representations sufficient for the individuation of visual words. In this late window, the lmFG became nearly insensitive to orthographic similarity as shown by similar classification accuracy for word pairs that differed by one letter compared with word pairs that were completely orthographically different (18). This kind of unique encoding of words is required to permit the individuation of visual words, a necessary step in word recognition (see Table 1 for summary). The time window in which this individuation signal is seen suggests that interactions with other brain regions transform the orthographic representation within the lmFG in support of word recognition. Such interactivity could function to integrate the orthographic, phonological, and semantic knowledge that together uniquely identifies a written word (23). Lack of spatiotemporal resolution to detect dynamic changes in lmFG coding of orthographic stimuli using fMRI may help to explain competing evidence for and against the visual word form hypothesis in the literature (5, 6).

The dynamic shift in the specificity of orthographic representation in the lmFG has a very similar time course as the coarse-to-fine processing shown in face-sensitive regions of the human fusiform (33). Considering that only an gist-level representation is available until ∼250 ms, and that saccade planning and execution generally occur within 200–250 ms during natural reading (34), the gist-to-individuated word-processing dynamic has important implications for neurobiological theories of reading; it suggests that when visual word form knowledge first makes contact with the language system, it is in the form of gist-level information that is insufficient to distinguish between visually similar alternatives. The identification of the early gist-level representation is consistent with evidence that readers are vulnerable to making errors in word individuation during natural reading, but contextual constraints are normally sufficient to avoid misinterpretations (35).

Table 1. Summary of electrophysiological results in early and late time windows

	Word category sensitivity		Bigram frequency sensitivity		Word individuation	
Patient number	Early	Late	Early	Late	Early	Late
P1	++	+	++	-	-	++
P2	++	+				
P3	++	+			-	++
P4	++	+	++	-	-	++

In other words, in most cases, accurate individuation is achieved through continued processing that likely involves mutually constraining orthographic, phonological, semantic, and contextual information, resulting in a more precise individuated word representation.

Another notable pattern in the gist-to-individuation temporal dynamic is that during the later time window when individuation is significant (~300–500 ms; Fig. 5), we found that the power to detect category-level word selectivity (i.e., words vs. bodies and scrambled images; Fig. 2), which arguably only requires gist-level discrimination, weakened and the event-related potential (ERP) response waned. This result is also consistent with a temporal selectivity pattern described for faces (33). One potential explanation for this selectivity and power shift could be that individuation is achieved by relatively few neurons (sparse coding) (36). Sparse coding would imply that relatively few word-sensitive neurons were active, and that the summed approximate word-related activity in this time period therefore would be weak. However, the neurons that were active encode for more precise word information, which would explain the significant word individuation reported here.

The mechanism underlying the representational shift from gist to individuation could have implications for models of reading disorders, such as dyslexia, where visual word identification is impaired (37). Indeed, the effects of lmFG stimulation, especially slower reading times, are suggestive of acquired (14) and developmental reading pathologies (38), which have been linked to dysfunction of lmFG (39). The extent to which individual word reading may be impaired by excess noise in the visual word form system, or the inadequate ability to contextually constrain noisy input into the language system, is for future research to untangle.

In summary, our results provide strong evidence that the lmFG is involved in at least two temporally distinguishable processing stages: an early stage that allows for category-level word decoding and gist-level representation organized by orthographic similarity, and a later stage supporting precise word individuation. An unanswered question is how the representation in the lmFG transitions between stages in these local neural populations and how interactions between areas involved in reading may govern these transitions. Taken together, the current results suggest a model in which lmFG contributes to multiple levels of orthographic representation via a dynamic shift in the computational analysis of different aspects of word information.

Methods

Subjects. Four patients (two males, ages 25–45) undergoing surgical treatment for medicine-resistant epilepsy participated in the experiments. The patients gave written informed consent to participate in this study, under a protocol approved by the University of Pittsburgh Medical Center Institutional Review Board. See *SI Methods* for demographic and clinical information about each participant.

Experimental Paradigm. The experiment paradigm and the data preprocessing method were similar to those described previously by Ghuman et al. (33). Paradigms were programmed in MATLAB using Psychtoolbox and custom-written code. All stimuli for the Category Localizer, Covert Naming, Word Individuation, and Stimulation were presented on a 22-inch LCD computer screen placed ~2 m from the participant's head at the center of the screen (~10 × 10° of visual angle). All stimuli for P1–P3 were identical. Due to a considerable delay in testing, the covert naming and word individuation stimuli were modified and updated for P4 to address additional questions beyond the scope of the current study. However, the critical characteristics of the stimuli and contrasts in the analyses remain consistent across all four patients. The category localizer was identical for all patients.

Category Localizer.
Stimuli. In the localizer experiment, 90 different images from three categories were used, with 30 images of bodies (50% male), 30 images of words, and 30 phase-scrambled images. Phase-scrambled images were created in MATLAB by taking the 2D Fourier transform of the image, extracting the phase, adding random phases, recombining the phase and amplitude, and taking the 2D inverse Fourier transform.
Design and procedure. In the category localizer, each image was presented for 900 ms with 900-ms intertrial intervals, during which a fixation cross was presented at the center of the screen. There were two consecutive blocks in a session. Each block consisted of all 180 images with a random presenting order. At random, one-third of the time an image would be repeated, which yielded a total of 480 trials in one recording session. The participant was instructed to press a button on a button box when an image was repeated (one-back task).

Electrical Brain Stimulation.
Stimuli. The stimuli used during electrode stimulation for P1 included 60 seven-letter words with 11.35 (10.60–13.67) mean log frequency, determined by the HAL Study used in the English Lexicon project (elexicon.wustl.edu/); single letters; and 13 famous faces that were familiar and nameable by P1. Stimuli were presented repeatedly during the session, starting with low-stimulation trials. Thus, stimuli presented during high-stimulation trials were likely to have been seen previously. The stimuli used during electrode stimulation for P2 included 46 seven-letter words with 10.93 (10.02–13.13) mean log frequency, and black-and-white pictures of common objects and animals. The 46 words that were presented during stimulation trials were out of a set of 155 words total that did not repeat.
Design and procedure. Electrical current during stimulation passed between adjacent electrode pairs (e.g., 1 and 2; 3 and 4; etc.). During the stimulation session presurgery, stimulation (1–10 mA, peak-to-peak amplitude, which is the distance between the negative and positive square waves delivered to the two contacts, i.e., this is 2× the amplitude of the square waves) was alternatingly applied with sham stimulation, whereas P1 and P2 overtly named words (P1 and P2), letters (P1), famous faces (P1), and pictures (P2). Each stimulus trial began with a beep, followed by 750 ms of fixation and then the stimulus. The stimulus remained on the screen until it was named, after which an experimenter manually advanced to the next item. Naming times were computed by calculating the time between the beep and the response (minus 750 ms). Only trials in which the electrode stimulation overlapped with the first 500 ms of stimulus presentation were included in further statistical analyses. T-tests comparing high- and low-stimulation trials were computed assuming unequal variances and df adjusted based on Levene's test for equality of variances.

Covert Naming: Sensitivity to Bigram Frequency.
Stimuli. In the covert word-naming experiment, words with nonoverlapping high- and low-bigram frequency (70 each for P1, 40 each for P4), controlled for lexical frequency, were used as visual stimuli.
Design and procedure. In the covert word-naming experiment, each word was presented once, in a random order, for 3,000 ms with 1,000-ms intertrial interval during which a fixation cross was presented at the center of the screen. The patient was instructed to press a button the moment when he began to covertly name the word to himself to ensure phonological encoding of each word and to avoid potential movement artifacts that could result from overt articulation.

Word Individuation.
Stimuli. In the word individuation experiment, 20 different English words, with word length ranging from two to five, were used as visual stimuli. Similar word pairs differed by one letter, and different word pairs did not share any letters. All comparisons were made within the same word length.
Design and procedure. In the word individuation experiment, each image was presented for 900 ms with 900-ms intertrial intervals, during which time a fixation cross was presented at the center of the screen. There were 24 consecutive blocks within a session. Each block consisted of all of the 20 words with a random order. At random, one-sixth of the time an image would be repeated, which yielded a total of 560 trials in one session. The patient was instructed to press a button on a button box when an image was repeated.

Multivariate Temporal Pattern Analysis. Considering that the size of the training set was smaller than the data dimensionality, a low-variance classifier (specifically, Gaussian naïve Bayes) was used. Principle component analysis (PCA) and linear discriminant analysis (LDA) were used to lower the dimensions in the case of multiway categorical classifications. However, we found the dimensionality reduction method was not plausible in the pairwise words classification case, because the smaller number of trials made the estimation of covariance unreliable. For all classification analyses, the Gaussian naïve Bayes classifier was trained based on the data from each time point of 100-ms windows from single trials in the training set (the time course pattern from

100 ms of single-trial potentials) and was used to label the condition of the corresponding data from that time window from the testing trial. The classification accuracy was estimated by counting the correctly labeled trials. This procedure was then repeated for all time windows slid with 10-ms steps between −100 and ∼600 ms relative to the presentation of the stimuli.

For the multiway categorical classifications with K categories (here, $K = 2$ or 3), the classification accuracy was estimated through nested leave-P-out cross-validation. In the first level of cross-validation, single-trial potentials were first split into training (80% of the trials) and testing set (20% of the trials) randomly. For each random split, PCA was trained based on the training set to lower the dimensionality down to P. Then, LDA was used to project the data into $K - 1$ dimensional space. Finally, a Gaussian naïve Bayes classifier was trained based on the projected training set. The selection of the model parameter P was achieved by finding the P that gave greatest d' for Bayes classification based on an additional level of random subsampling validation with 50 repeats using only the training set. After training, true positive and false alarm rates of the target condition were calculated across all of the test trials. The d' was calculated as $d' = Z(\text{true positive rate}) - Z(\text{false alarm rate})$, where Z is the inverse of the Gaussian cumulative distribution function. The random split was repeated 200 times, and the classification accuracy was estimated by averaging across results from these 200 random splits.

For the pairwise classification in the word individuation task, the pairwise classification accuracy was estimated through leave-one-out cross-validation. Specifically, for each pair of words, each trial was left out in turn as the testing trial, with the remaining trials used for the training set. Finally, the overall pairwise classification accuracy was estimated through averaging across all 190 word pairs. The classification accuracy for each specifically controlled condition was estimated by averaging the corresponding word pairs.

See *SI Methods* for details regarding statistical testing of classification accuracy.

ACKNOWLEDGMENTS. We thank the patients and their families for their time and participation; the epilepsy monitoring unit staff, Cheryl Plummer, Gena Ghearing, and administration for assistance and cooperation with our research; Breana Gallagher for assistance with coding and analysis; Ellyanna Kessler, Roma Konecky, Nicolas Brunet, and Witold Lipski for assistance with data collection; Marlene Behrmann for assistance and access to stimuli for the letter-length neuropsychological test; and Daphne Bavelier and Charles Perfetti for helpful comments and feedback on this work. This work was supported by National Institute of Neurological Disorders and Stroke Award T32NS086749 (to E.A.H.), National Institute on Drug Abuse Awards R90DA023426 and R90DA023420 (to Y.L.), Eunice Kennedy Shriver National Institute of Child Health and Human Development Award R01HD060388 (to J.A.F.), and National Institute of Mental Health Award NIH R01MH107797 (to A.S.G.).

1. Bub DN, Arguin M, Lecours AR (1993) Jules Dejerine and his interpretation of pure alexia. *Brain Lang* 45(4):531–559.
2. Wernicke C (1977) Der aphasischer symptomenkomplex: Eine psychologische studie auf anatomischer basis. Wernicke's Works on Aphasia: A Sourcebook and Review, ed Eggert GH (Mouton, The Hague), pp 91–145.
3. Warrington EK, Shallice T (1980) Word-form dyslexia. *Brain* 103(1):99–112.
4. Dehaene S, Le Clec'H G, Poline JB, Le Bihan D, Cohen L (2002) The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *Neuroreport* 13(3):321–325.
5. Dehaene S, Cohen L (2011) The unique role of the visual word form area in reading. *Trends Cogn Sci* 15(6):254–262.
6. Price CJ, Devlin JT (2011) The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci* 15(6):246–253.
7. Dehaene S, et al. (2010) How learning to read changes the cortical networks for vision and language. *Science* 330(6009):1359–1364.
8. Schlaggar BL, McCandliss BD (2007) Development of neural systems for reading. *Annu Rev Neurosci* 30:475–503.
9. Brem S, et al. (2010) Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc Natl Acad Sci USA* 107(17):7939–7944.
10. Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2011) The development of cortical sensitivity to visual word forms. *J Cogn Neurosci* 23(9):2387–2399.
11. Xue G, Poldrack RA (2007) The neural substrates of visual perceptual learning of words: Implications for the visual word form area hypothesis. *J Cogn Neurosci* 19(10): 1643–1655.
12. Glezer LS, Kim J, Rule J, Jiang X, Riesenhuber M (2015) Adding words to the brain's visual dictionary: Novel word learning selectively sharpens orthographic representations in the VWFA. *J Neurosci* 35(12):4965–4972.
13. Gaillard R, et al. (2006) Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50(2):191–204.
14. Behrmann M, Shallice T (1995) Pure alexia: A nonspatial visual disorder affecting letter activation. *Cogn Neuropsychol* 12(4):409–454.
15. McCandliss BD, Cohen L, Dehaene S (2003) The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn Sci* 7(7):293–299.
16. Binder JR, Medler DA, Westbury CF, Liebenthal E, Buchanan L (2006) Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33(2): 739–748.
17. Cohen L, et al. (2002) Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain* 125(Pt 5):1054–1069.
18. Glezer LS, Jiang X, Riesenhuber M (2009) Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron* 62(2):199–204.
19. Farah MJ, Wallace MA (1991) Pure alexia as a visual impairment: A reconsideration. *Cogn Neuropsychol* 8(3-4):313–334.
20. Price CJ, Devlin JT (2003) The myth of the visual word form area. *Neuroimage* 19(3): 473–481.
21. Vinckier F, et al. (2007) Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55(1):143–156.
22. Wandell BA (2011) The neurobiological basis of seeing words. *Ann N Y Acad Sci* 1224(1):63–80.
23. Whaley ML, Kadipasaoglu CM, Cox SJ, Tandon N (2016) Modulation of orthographic decoding by frontal cortex. *J Neurosci* 36(4):1173–1184.
24. Baeck A, Kravitz D, Baker C, Op de Beeck HP (2015) Influence of lexical status and orthographic similarity on the multi-voxel response of the visual word form area. *Neuroimage* 111:321–328.
25. Maurer U, Brandeis D, McCandliss BD (2005) Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behav Brain Funct* 1(1):13.
26. Nobre AC, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. *Nature* 372(6503):260–263.
27. Hamamé CM, et al. (2013) Dejerine's reading area revisited with intracranial EEG: Selective responses to letter strings. *Neurology* 80(6):602–603.
28. Hamamé CM, et al. (2014) Functional selectivity in the human occipitotemporal cortex during natural vision: Evidence from combined intracranial EEG and eye-tracking. *Neuroimage* 95:276–286.
29. Mani J, et al. (2008) Evidence for a basal temporal visual language center: Cortical stimulation producing pure alexia. *Neurology* 71(20):1621–1627.
30. Parvizi J, et al. (2012) Electrical stimulation of human fusiform face-selective regions distorts face perception. *J Neurosci* 32(43):14915–14920.
31. Tanji K, Suzuki K, Delorme A, Shamoto H, Nakasato N (2005) High-frequency γ-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. *J Neurosci* 25(13):3287–3293.
32. Duncan KJ, Pattamadilok C, Devlin JT (2010) Investigating occipito-temporal contributions to reading with TMS. *J Cogn Neurosci* 22(4):739–750.
33. Ghuman AS, et al. (2014) Dynamic encoding of face information in the human fusiform gyrus. *Nat Commun* 5:5672.
34. Reichle ED, Pollatsek A, Fisher DL, Rayner K (1998) Toward a model of eye movement control in reading. *Psychol Rev* 105(1):125–157.
35. Levy R, Bicknell K, Slattery T, Rayner K (2009) Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proc Natl Acad Sci USA* 106(50):21086–21090.
36. Young MP, Yamane S (1992) Sparse population coding of faces in the inferotemporal cortex. *Science* 256(5061):1327–1331.
37. Bruck M (1990) Word-recognition skills of adults with childhood diagnoses of dyslexia. *Dev Psychol* 26(3):439.
38. Bowers PG, Wolf M (1993) Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Read Writ* 5(1):69–85.
39. Martin A, Kronbichler M, Richlan F (April 7, 2016) Dyslexic brain activation abnormalities in deep and shallow orthographies: A meta-analysis of 28 functional neuroimaging studies. *Hum Brain Mapp*, 10.1002/hbm.23202.
40. Torgesen JK, Wagner R, Rashotte C (1999) *TOWRE–2 Test of Word Reading Efficiency* (Pro-Ed, Austin, TX).
41. Wagner RK, Torgesen JK, Rashotte CA (1999) *Comprehensive Test of Phonological Processing* (Pro-Ed, Austin, TX).
42. Glezer LS, Riesenhuber M (2013) Individual variability in location impacts orthographic selectivity in the "visual word form area". *J Neurosci* 33(27):11221–11226.
43. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.
44. Nestor A, Behrmann M, Plaut DC (2013) The neural basis of visual word form processing: A multivariate investigation. *Cereb Cortex* 23(7):1673–1684.
45. Miller KJ, Schalk G, Hermes D, Ojemann JG, Rao RP (2016) Spontaneous decoding of the timing and content of human object perception from cortical surface recordings reveals complementary information in the event-related potential and broadband spectral change. *PLOS Comput Biol* 12(1):e1004660.
46. Shum J, et al. (2013) A brain area for visual numerals. *J Neurosci* 33(16):6709–6715.