Orthographic contamination of Broca's area

Marie Montant[1]*, Daniele Schön[2], Jean-Luc Anton[3] and Johannes C. Ziegler[1]

[1] Laboratoire de Psychology Cognitive, CNRS, Aix-Marseille University, Marseille, France
[2] CNRS, Institut de Neurosciences Cognitives de la Méditerranée, Aix-Marseille University, Marseille, France
[3] Centre d'IRM Fonctionnelle Cérébrale de Marseille, IFR 131, Marseille, France

Strong evidence has accumulated over the past years suggesting that orthography plays a role in spoken language processing. It is still unclear, however, whether the influence of orthography on spoken language results from a co-activation of posterior brain areas dedicated to low-level orthographic processing or whether it results from orthographic restructuring of phonological representations located in the anterior perisylvian speech network itself. To test these hypotheses, we ran a fMRI study that tapped orthographic processing in the visual and auditory modalities. As a marker for orthographic processing, we used the orthographic decision task in the visual modality and the orthographic consistency effect in the auditory modality. Results showed no specific orthographic activation neither for the visual nor the auditory modality in left posterior occipito-temporal brain areas that are thought to host the visual word form system. In contrast, specific orthographic activation was found both for the visual and auditory modalities at anterior sites belonging to the perisylvian region: the left dorsal–anterior insula and the left inferior frontal gyrus. These results are in favor of the restructuring hypothesis according to which learning to read acts like a "virus" that permanently contaminates the spoken language system.

Keywords: visual word recognition, speech perception, orthographic consistency, inferior frontal gyrus, insula, visual word form system, visual word form area

"In literate adults, orthography is important in speech recognition just as phonology is important in reading"

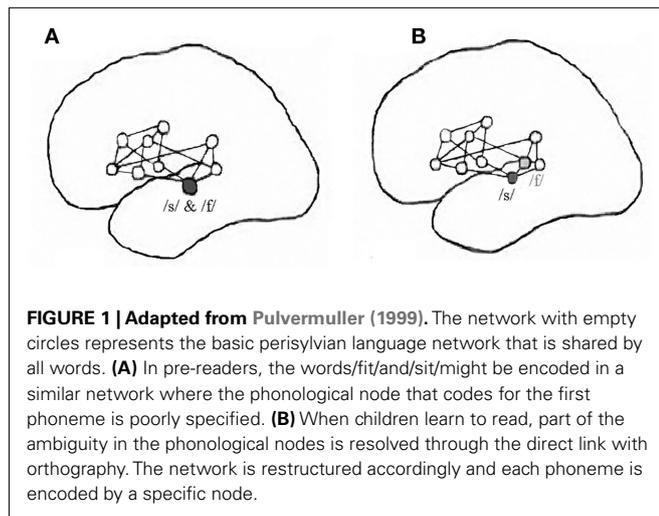*(Taft et al., 2008, p.360).*

INTRODUCTION

Children learn to speak before they learn to read and write. The acquisition of spoken words is based on the development of phonological skills and the mapping of speech sounds onto meaning (e.g., Curtin and Werker, 2007). Sound and meaning of words are primarily represented in the brain in the left cerebral hemisphere, namely in left perisylvian structures including Wernicke's and Broca's area (e.g., Hickok and Poeppel, 2007). In adults, the basic cerebral network for words includes the superior temporal gyrus (STG), the superior temporal sulcus (STS), the supramarginal gyrus (SMG), the inferior frontal gyrus (IFG), and the more dorsal/anterior premotor cortex (e.g., Pulvermuller et al., 2009; see **Figure 1** below).

In the process of learning to read and write, one can imagine at least two possible ways that orthography might be implemented in the brain. First, occipito-temporal structures that have been shaped to process visual objects in the course of evolution could be recruited to process written language. This is the *recycling hypothesis* of the left fusiform gyrus proposed by the Dehaene and Cohen (2007). According to this hypothesis, the left fusiform gyrus is in charge of the orthographic form of words and hosts the visual word form system (VWFS; Carr, 1986), also known in the brain imaging literature as visual word form area (VWFA; McCandliss et al., 2003). This system processes all types of letter strings, from single letters to real words, following a postero-anterior gradient

(Vinckier et al., 2007) with letter strings of higher bigram frequency being processed in a more anterior part of the fusiform gyrus than letter strings of lower bigram frequency.

The second way that learning to read could be implemented in the brain is to restructure the already existing language (speech) network situated in the perisylvian region. More precisely, this means that the phonological and semantic representations of spoken words are remodeled to include orthographic information. This second hypothesis is appealing for two reasons. First, it does not rely on the building of an entirely new cerebral network for reading. Second, orthography might help to reduce the ambiguity that is inherent in the speech signal. Developmental studies indeed suggest that learning to read improves the quality of the phonological representations (Goswami et al., 2005; Ventura et al., 2007; Ziegler and Muneaux, 2007). For instance, the spoken words/sit/and/fit/might be difficult to discriminate during early language development because they differ only by one phonetic feature, namely the place of articulation of the first phoneme. During the process of learning to read, information from the visual system (as well as somatosensory and motor cues resulting from handwriting movements) might help to disambiguate the auditory signal and, as a consequence, create, or consolidate distinct phonological representations for these two words. Restructuring implies changes in the connectivity of the speech network and a rearrangement of the nodes that represent these words (see **Figure 1**).

There are indeed several brain imaging studies suggesting that learning to read re-shapes the cerebral network for speech perception (Castro-Caldas et al., 1998; Carreiras et al., 2009; Dehaene

FIGURE 1 | Adapted from Pulvermuller (1999). The network with empty circles represents the basic perisylvian language network that is shared by all words. **(A)** In pre-readers, the words /fit/ and /sit/ might be encoded in a similar network where the phonological node that codes for the first phoneme is poorly specified. **(B)** When children learn to read, part of the ambiguity in the phonological nodes is resolved through the direct link with orthography. The network is restructured accordingly and each phoneme is encoded by a specific node.

et al., 2010). For example, in an auditory word repetition task that does not require orthographic knowledge, illiterate people do not activate the speech network hosted in left temporal brain regions the way literate people do (Castro-Caldas et al., 1998; Petersson et al., 2000). Similarly, literates show greater activation than illiterates in the planum temporale, a region typically associated with phoneme processing (Dehaene et al., 2010). Finally, literates exhibit more white matter than illiterates in brain regions typically associated with phonological processing, such as the bilateral angular, dorsal occipital, middle temporal, left supramarginal, and STG (Carreiras et al., 2009).

In the psycholinguistic literature, a growing number of studies have shown that speech perception is automatically influenced by orthographic knowledge even when participants are totally unaware of any orthographic manipulation (e.g., Taft et al., 2008). One of the clearest demonstrations was provided by Ziegler and Ferrand (1998). They manipulated the consistency with which phonology mapped onto spelling in an auditory lexical decision task. Inconsistent words, that is, words whose rhymes could be spelled in multiple ways (e.g., /-ip/ may be spelled "-eap" or "-eep"), produced slower correct "yes" responses and more errors than consistent words (e.g., "duck"; /-uk/ may only be spelled "-uck"). This orthographic consistency effect has been replicated in different languages (Ventura et al., 2004; Pattamadilok et al., 2007; Ziegler et al., 2008), different tasks (Pattamadilok et al., 2009; Peereman et al., 2009), and with different orthographic manipulations (e.g., Ziegler et al., 2003, 2004).

An outstanding question concerns the locus of the orthography effect on spoken language. Does the influence of orthography on spoken language result from a co-activation of the VWFS, dedicated to low-level orthographic processing and located in posterior brain areas (i.e., left fusiform gyrus), or does it result from orthographic restructuring of phonological representations located in the anterior perisylvian speech network itself? The first hypothesis would be in line with the "recycling" view of reading development (e.g., Dehaene and Cohen, 2007), in which orthographic processing happens in occipito-temporal brain structures that, prior to reading, were in charge of processing visual objects and faces. The

second hypothesis would be in line with the "restructuring" view of reading development (Ehri, 1992; Ziegler and Goswami, 2005; Goswami and Ziegler, 2006), in which orthography is *amalgamated* within a widely distributed spoken language system (e.g., IFG, STG, and SMG). According to the restructuring hypothesis, the VWFS would merely provide the visuo-orthographic "entry gate" to the spoken language system but orthographic information could be represented well beyond the left fusiform gyrus.

Behavioral data concerning the orthographic consistency effect on spoken language (e.g., Ziegler and Ferrand, 1998) are not able to decide between these two hypotheses as the same pattern of results (e.g., slower responses and more errors to inconsistent than to consistent words) can either result from on-line activation of distant posterior structures (left fusiform gyrus) or from the encoding of orthographic knowledge in the anterior perisylvian speech network itself. However, two recent studies are in favor of the restructuring hypothesis. First, Perre et al. (2009) showed that the cortical generators of the orthographic consistency effect obtained in ERPs at 350 ms were localized in a left temporo-parietal area, including parts of the SMG (BA40), the posterior STG (BA22) and the inferior parietal lobule (BA40). Second, Pattamadilok et al. (2010) showed that transcranial magnetic stimulation of the left SMG (but not the left fusiform gyrus) removed the orthographic consistency effect in an auditory lexical decision task.

The purpose of the present study was to shed further light on the locus of the orthographic consistency effect and to provide *direct* evidence in favor of the restructuring or the recycling hypothesis. To this end, we ran a fMRI study that tapped orthographic processing in the visual and auditory modalities. In the visual modality, we used the orthographic decision task as a marker for orthographic processing. That is, on a given trial, we presented a pseudohomophone together with its base word (e.g., BRANE–BRAIN) and asked participants to decide which of the two was spelled like the real word. Given that a word and its pseudohomophone have the same phonology, participants must use *lexical* orthographic knowledge to make their decision. Note that this task taps higher-level orthographic information (i.e., lexical orthography) than the tasks that are typically used to tap the VWFS (e.g., passive viewing of letter strings). In the auditory modality, we used the orthographic consistency effect as a marker for orthographic processing. That is, we ran a lexical decision task with orthographically consistent and inconsistent words (see above).

The core hypotheses were the following: if the orthographic consistency effect in spoken language resulted from co-activation of the VWFS, as predicted by the recycling view of reading development, then we should obtain orthographic effects in posterior brain areas (i.e., left fusiform gyrus). In contrast, if the orthographic consistency effect resulted from orthographic restructuring of phonological representations, we should obtain orthographic effects in the anterior perisylvian speech network itself. Furthermore, if the same kind of orthographic codes were involved both in visual and auditory word recognition, then we should obtain orthography-related activation of the same brain regions in the visual and auditory modalities. Shared activation across visual and auditory tasks in posterior brain regions would support the recycling hypothesis, while shared activation in anterior

(perisylvian) regions would be in favor of the restructuring hypothesis.

MATERIALS AND METHODS
PARTICIPANTS
Fourteen students at the University of Provence (nine women; mean age = 22; range = 18–27) participated in this study. All were right handed and French native speakers. They reported normal or corrected-to-normal vision, normal hearing, and no history of neurological problem. Participants gave written consent and were paid for their participation.

TASKS AND STIMULI
Visual modality
We used two tasks in the visual modality: an orthographic decision and a visual control task. In both tasks, a fixation cross was presented at the center of the screen for 2 s. The cross was then replaced by a pair of lower-case letter-strings that appeared simultaneously to the left and the right of the center for 1 s (see **Figure 2**). Each pair of letter strings consisted of a French word (e.g., *entier*) and a matched pseudohomophone (i.e., same phonology as the base-word, different spelling, e.g., *antier*). The stimulus material was composed of 100 high-frequency words (frequency: 187 ± 25 occurrences per million; length: four to six-letters) selected from Lexique 3[1] and 100 pseudohomophones that were created from the selected words by changing one letter at any position in the string.

In the orthographic decision task, participants were asked to decide as rapidly as possible which of the two letter strings of each pair was spelled correctly (i.e., which one was a real word). They gave their response with the right hand, by using a MRI compatible button box: left button for the stimulus to the left, right button for the stimulus to the right. The left/right position on the screen

[1]http://www.lexique.org/

of the word and pseudohomophone was counter-balanced across trials and participants.

In the visual control task, we used the same pairs of stimuli as described above, except that there was an upper-case letter in the center position of one of the two letter strings. The participants were asked to decide as rapidly as possible which of the letter-strings – left or right – contained an upper-case letter. They gave their response using the same procedure as described above for the orthographic decision task. The left/right position of the upper-case letter and the type of letter-string (word/pseudohomophone) was counter-balanced across trials and participants.

In both tasks, the letter-strings disappeared with the participant's button press. After each trial, there was an inter-stimulus interval (ISI) of 2 s, during which the fixation cross of the next trial was presented.
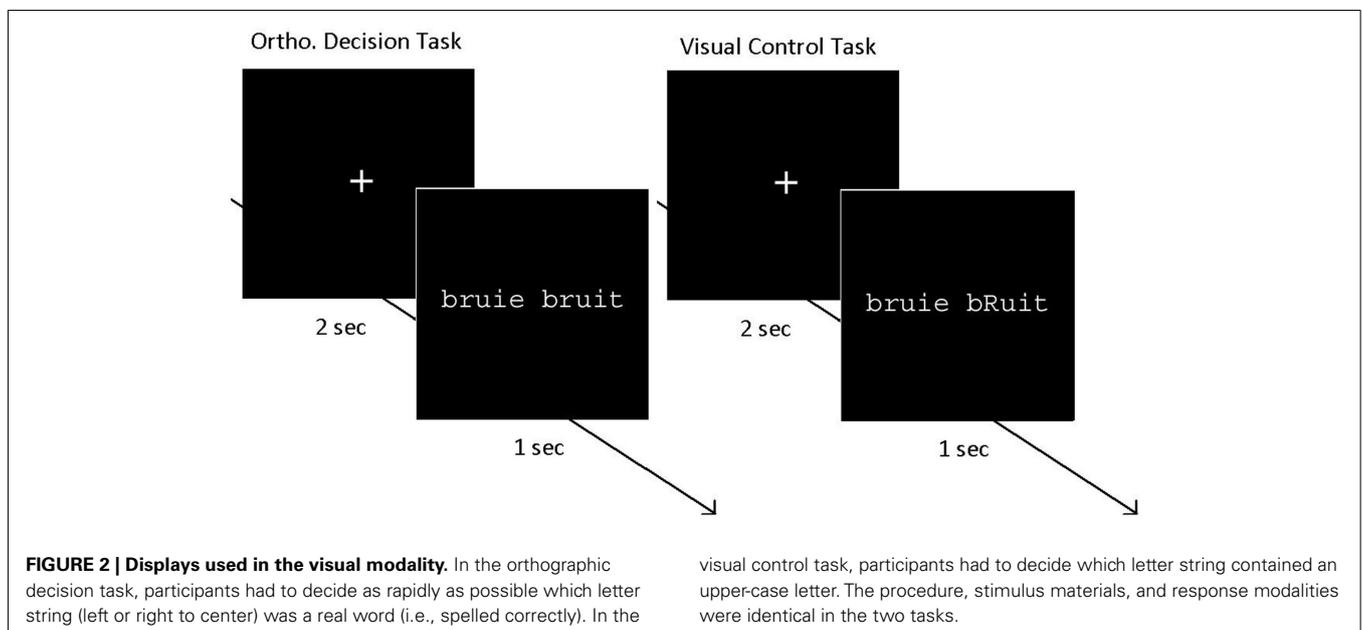
We used a block design in the visual modality. The orthographic decision and visual control tasks alternated every 10 trials (10 trials = 1 block). At the beginning of each block, a letter was presented for 1 s that indicated the nature of the block, the letter "M" cued the orthographic decision task, whereas the letter "L" cued the visual control task. The presentation duration of the fixation cross at the beginning of each block varied from 1 to 3 s to avoid the convolution of the BOLD signal with task switching. The order of the blocks was counter-balanced across participants.

The participants performed 10 blocks of 10 trials each, for a total duration of 12 min and 42 s. A short pause was inserted halfway through the experiment. Visual stimuli were projected onto a screen which was viewed by participants through a mirror positioned above their eyes.
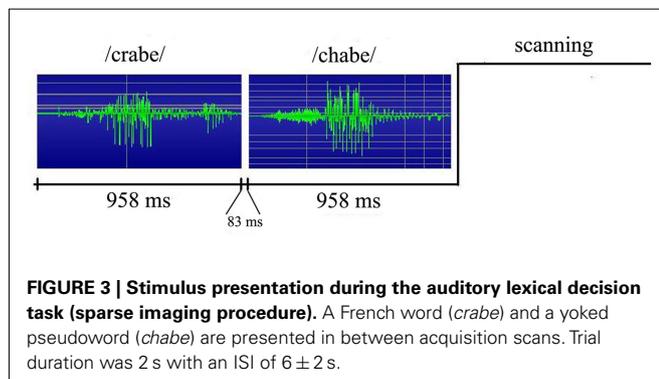
Auditory modality
Two tasks were used in the auditory modality: a lexical decision task and an auditory control task. In the lexical decision task, on a given trial, participants listened to a word and its matched pseudo-word (e.g., *crabe*, *chabe*), which were presented one after the other

FIGURE 2 | Displays used in the visual modality. In the orthographic decision task, participants had to decide as rapidly as possible which letter string (left or right to center) was a real word (i.e., spelled correctly). In the visual control task, participants had to decide which letter string contained an upper-case letter. The procedure, stimulus materials, and response modalities were identical in the two tasks.

in a randomized order through earphones. One of these items was pronounced by a male speaker while the other was pronounced by a female speaker. We used different voices in the auditory lexical decision task to match the acoustic conditions of the auditory control task (see below). Participants had to decide as rapidly as possible which of the two items – the first or the second – was a real word, independent of the voice of the speaker. Participants gave their response by using a MRI compatible button box (left button for the first item, right button for the second item). The stimulus material used in this task was composed of 80 monosyllabic French words selected from Lexique 3 (see text footnote 1): 40 consistent words (e.g., *stage*) and 40 inconsistent words (e.g, *faim*). Consistent and inconsistent words were matched on the frequency of occurrence, orthographic and phonological uniqueness point, number of phonemes, number of letters, number of homographs, number of homophones, and number of orthographic and phonological neighbors (all $ps > 0.1$). On the basis of the selected words, 80 matched pseudowords were created by changing the onset of the base words. Altogether, there were a total of 80 trials, half of which contained a consistent word and its yoked pseudoword while the other half contained an inconsistent word and its yoked pseudoword. The order of consistent and inconsistent trials was randomized. The order of presentation (first vs. second) of the word and the pseudoword in each trial and the voice of the speaker (male vs. female) that pronounced each of them were randomized as well. The duration of each word and pseudoword was 958 ms. Stimuli were slightly expanded or compressed in order to obtain a duration of 958 ms. There was a 83.3 ms interval between the two items on a given trial, such that the total duration of a trial was 2 s (see **Figure 3**).

In the auditory control task, on a given trial, participants were listening to pairs of vowels presented sequentially through earphones. One vowel was pronounced by a male speaker, whereas the other was pronounced by a female speaker. Participants had to decide which one, the first of the second, was pronounced by a male speaker. Participants were asked to answer as rapidly as possible, using a MRI compatible button box. As stimulus material, we used the vowels "a," "e," "i," "o," and "u," recorded by both a male and a female French speaker. Each vowel was repeated 16 times in order to obtain 80 trials. Timing was identical to the one described above for the lexical decision task. The order of presentation of the vowels (female vs. male voice) in each trial was randomized across the experiment.

FIGURE 3 | Stimulus presentation during the auditory lexical decision task (sparse imaging procedure). A French word (*crabe*) and a yoked pseudoword (*chabe*) are presented in between acquisition scans. Trial duration was 2 s with an ISI of 6 ± 2 s.

The auditory experiments (lexical decision and control tasks) were conducted using sparse imaging sampling to avoid scanner noise interference with the presentation of the language material. As above (see **Figure 3**), auditory stimuli were presented in a silent period of 2 s and scanning took place in the 6 s ISI between trials. In order to record separately the BOLD response to consistent and inconsistent words in the lexical decision task, the ISI between two consecutive trials was 6 s on average (±2 s to allow deconvolution of the BOLD signal from the experimental design). The lexical decision and the control tasks alternated every 10 trials. At the beginning of each block of lexical decision, the upper-case letter "M" (for Word) appeared on the screen for 1 s, followed by a 3 s delay before the first auditory stimulus was presented in earphones. At the beginning of the control task, the upper-case letter "S" (for Gender) appeared on the screen for 1 s, followed by a 3 s delay. Each block was repeated eight times, for a total duration of 24 min, with a short pause halfway through the experiment. The order of the blocks was counter-balanced across participants. Auditory stimuli were presented via dedicated and MRI compatible headphones.

MRI ACQUISITION AND PREPROCESSING

Data acquisition was performed on a 3-T MEDSPEC 30/80 AVANCE imager (Bruker, Eittlingen, Germany) at the fMRI center of Marseille, France.

Functional images were acquired using a T2*-weighted gradient-echo planar sequence (TR = 3000 ms; FOV = 192 mm × 192 mm, 64 × 64 matrix). Whole brain volumes were acquired in 36 interleaved axial slices (3 mm thick, no gap) parallel to the inter-commissural plane (voxel size 3 mm × 3 mm × 3 mm). For each participant, T1-weighted anatomical images (MPRAGE sequence, voxel size 1 mm × 1 mm × 1 mm) were acquired following the fMRI scanning sessions.

The fMRI data were preprocessed and analyzed using SPM2 software (Welcome Institute of Cognitive Neurology, London, UK[2]). The first two volumes of each run were discarded in order to allow for signal equilibrium. Preprocessing comprised within-subject spatial and temporal realignment, spatial normalization of images to a template in standard space Montreal Neurological Institute (MNI), and a spatial smoothing using a 6 mm Gaussian kernel.

MRI ANALYSES

Visual modality

Statistical analyses were also performed with SPM2 toolbox using a general linear model employing a boxcar function convolved with a hemodynamic response function. High pass filtering (cut-off period equal to 128 s) was carried out to reduce scanner and physiological artifacts. Auto-regression was used to correct for serial correlations. A fixed effect analysis was first employed with a regressor for each experimental condition. Task instructions were added as a regressor-of-no-interest. Each contrast was then used in a random effect analysis (t-test) for the contrast of interest. The statistical threshold was set to $p < 0.001$ and to a cluster size of

---

[2]http://www.fil.ion.ucl.ac.uk/spm/

at least 10 voxels. Activated brain regions were labeled using MNI Space Utility (MSU) toolbox[3].

Auditory modality

Analyses were similar to those carried out in the visual modality except for the fact that conditions were modeled as events rather than epochs. A fixed effect analysis was first employed with a regressor for each experimental condition (consistent word, inconsistent word, female-voice, and male-voice). Contrasts of interest (lexical decision vs. auditory and consistent vs. inconsistent) were then used in a random effect analysis (t-test).

RESULTS

BEHAVIORAL RESULTS

Behavioral analyses were run on the data of 12 participants only (the data of 2 participants were lost in a computer crash).

In the visual modality, there was a main effect of the tasks on reaction times (RTs); $[t(11) = 7.2; p < 0.0001]$ and on errors $[(t(11) = 3.6; p < 0.005]$. Participants were more accurate and 232 ms faster in the visual control than in the orthographic decision task (mean RTs = 844 and 1076 ms, respectively).

In the auditory modality, there was a main effect of the task on both RTs and errors $[t(11) = 5.8, p < 0.001$ and $t(11) = 9.4, p < 0.001$, respectively]. Participants were more accurate and much faster in the control task than in the lexical decision task (mean RTs = 1478 and 2571 ms, respectively). In the auditory lexical decision (see **Figure 4**), there was a main effect of consistency on RTs $[F(1,11) = 11.48; p < 0.01]$, and no interaction between consistency and the order of presentation of the word and the pseudoword (first/second) in a trial $[F(1,11) = 2.35, p > 0.1]$. Responses to consistent words were 96 ms faster than responses to inconsistent words. The consistency effect on error rates was not statistically significant (15.6% errors for consistent words, 14.2% for inconsistent words; all $ps > 0.1$).
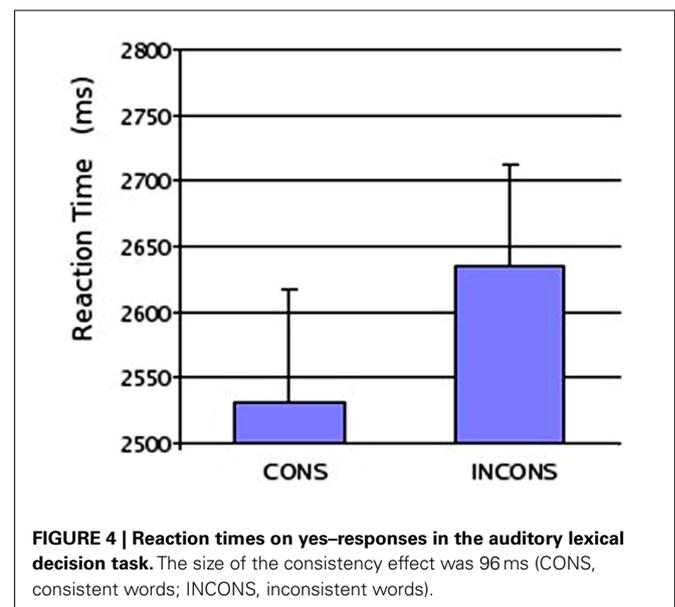
BRAIN IMAGING RESULTS

In the visual modality, we first identified the regions activated specifically in the visual orthographic decision task compared

[3]http://www.ibh.spb.ru/pet_lab/MSU/MSUMain.html

to the visual control task (see **Table 1**; **Figure 5**). In the posterior brain, we found strong bilateral activation in the region of the calcarine sulcus, extending to the cuneus, the lingual gyrus, the medial occipito-temporal cortex, and the cerebellum (see **Figure 5A,B**). The VWFA in the left fusiform gyrus was not significantly more activated in the orthographic than in the control condition. In the anterior brain, we observed left hemispheric activation of the IFG and anterior insula (see **Figure 5C**).

In the auditory modality, we first identified the regions activated specifically in the lexical decision compared to the auditory control task (see **Table 2** and **Figure 6**). As expected, there was strong activation of the perisylvian region. This included the left superior and middle temporal gyri, the left inferior parietal lobule, and a large part of the frontal lobe, mostly in the pre-central region, including the superior frontal gyrus and the IFG.

In order to identify more precisely the brain regions involved in orthographic processing *per se*, we conducted two additional analyses: an inclusive masking of visual and auditory word

FIGURE 4 | Reaction times on yes–responses in the auditory lexical decision task. The size of the consistency effect was 96 ms (CONS, consistent words; INCONS, inconsistent words).

Table 1 | Activated regions in the visual orthographic decision compared to visual control task.

Region			Z-score	Maximum peak coordinate (MNI)		
				x	y	z
Occipital	Med.	Calcarine sulcus	4.53	4	−86	6
	L	Cuneus/lingual gyrus	4.19	−22	−86	−2
	R	Cuneus/lingual gyrus	4.08	20	−80	2
Occipito-temporal	L	Medial occipital/temporal junction	4.42	−38	−58	10
Frontal	L	Anterior insula	3.94	−34	20	2
	L	IFG	3.66	−42	24	12
Subcortical	R	Cerebellum	4.99	6	−64	−8
	L	Midbrain	3.49	−10	−26	−10

Montreal Neurological Institute coordinate of the maximum peak activation for each cluster (>10 voxels) is provided together with Z-scores at the voxel level.
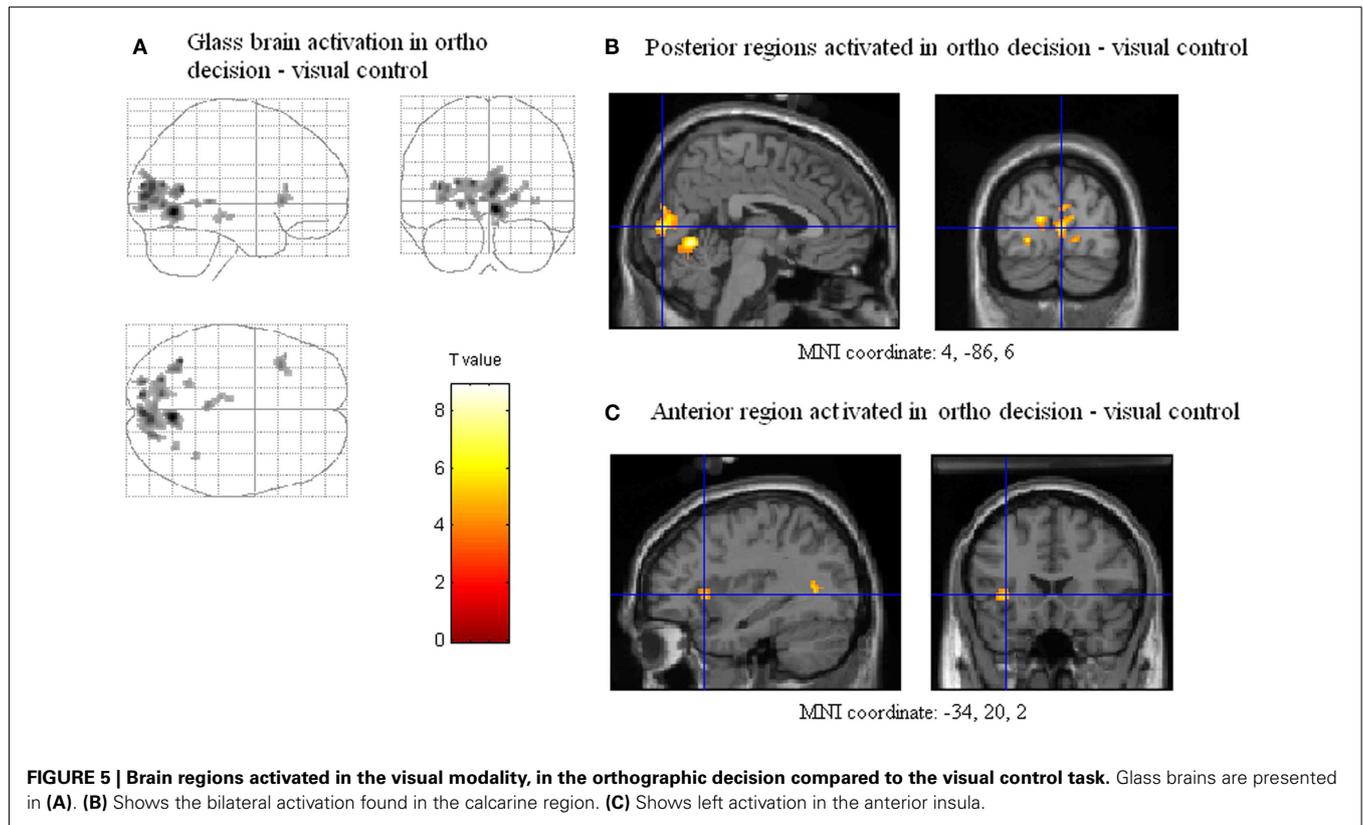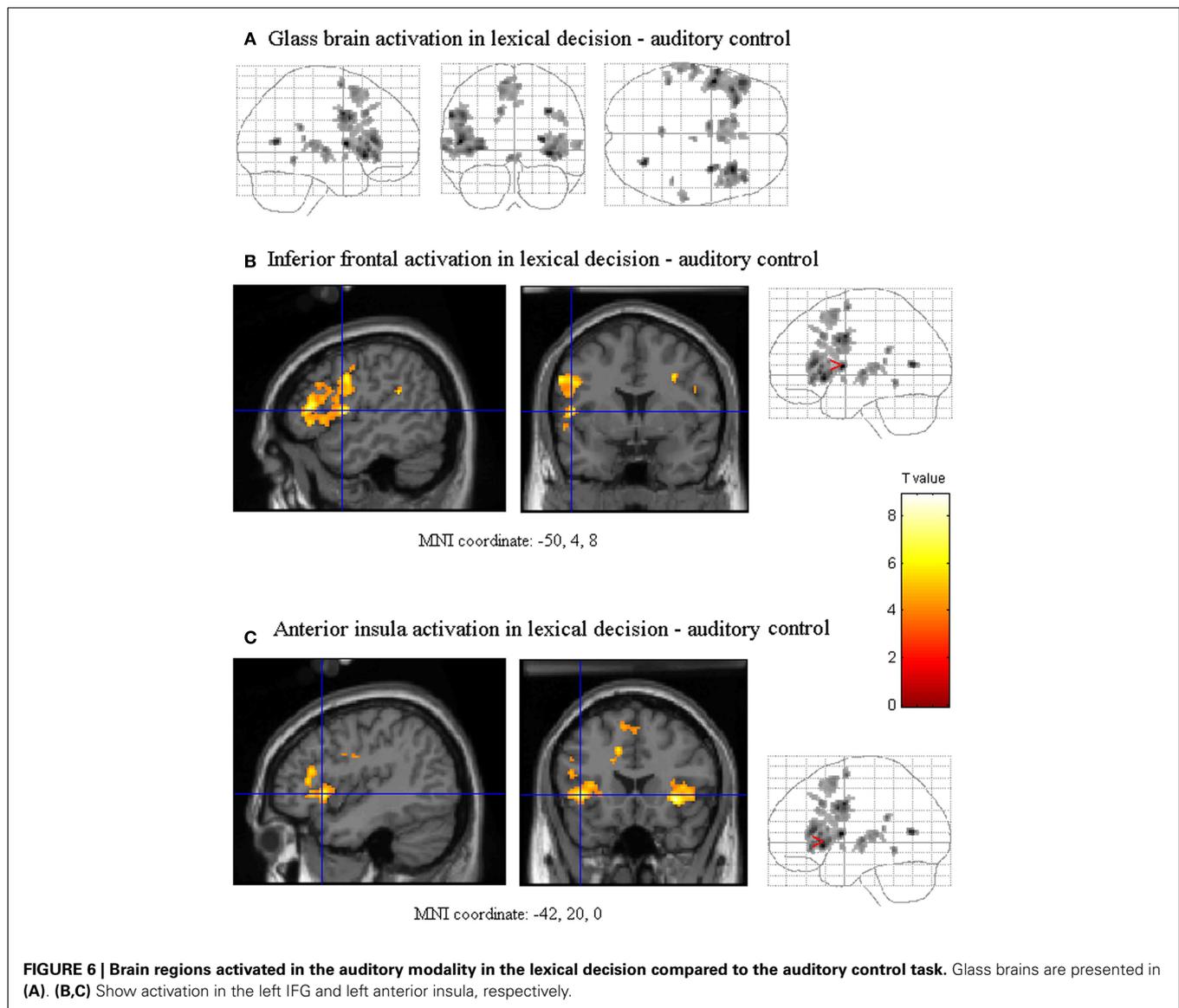
FIGURE 5 | Brain regions activated in the visual modality, in the orthographic decision compared to the visual control task. Glass brains are presented in **(A)**. **(B)** Shows the bilateral activation found in the calcarine region. **(C)** Shows left activation in the anterior insula.

Table 2 | Activated regions in the auditory lexical decision compared to auditory control task.

Region			Z-score voxel	Maximum peak coordinate (MNI)		
				x	Y	z
Temporal	L	Superior temporal gyrus	4.17	−66	−16	2
	L	Middle temporal gyrus	3.76	−56	−16	−4
	R	Superior temporal gyrus	3.82	62	−24	6
Parietal	L	Inferior parietal lobule	4.03	−50	−40	24
Frontal	L	Superior frontal gyrus	3.91	4	10	60
	L	Insula	4.54	−42	20	0
	L	IFG	4.96	−50	4	8
	L	Pre-central gyrus	4.60	−58	6	32
	L	Medial frontal gyrus	4.10	−8	−2	64
	L	Cingulate gyrus	4.03	−12	22	36
	R	IFG	4.19	44	32	−6
	R	Insula	4.78	34	20	−4
	R	Pre-central gyrus	4.64	32	0	36
Subcortical	R	Post-cingulate	4.61	26	−62	10
	L	Midbrain				
	L	Cerebellum	4.02	−4	−46	−8

Montreal Neurological Institute coordinate of the maximum peak activation for each cluster (>10 voxels) is provided together with Z-scores at the voxel level.

recognition [i.e., (ortho decision − visual control) inclusively masked by (lexical decision − auditory control); $p > 0.001$; cluster size: 10 voxels; see **Figure 7A**), and a contrast analysis between inconsistent and consistent word trials (orthographic consistency effect) in the auditory modality ($p < 0.001$; cluster size: 10 voxels; see **Figure 7B**). The inclusive masking analysis resulted in the activation of left insula (MNI coordinate: −34, 20, 2; Z-score at the voxel level: 3.94). The contrast analysis between

FIGURE 6 | Brain regions activated in the auditory modality in the lexical decision compared to the auditory control task. Glass brains are presented in **(A)**. **(B,C)** Show activation in the left IFG and left anterior insula, respectively.

inconsistent vs. consistent trials resulted in the activation of the left IFG (MNI coordinate: −54, 20, 10; Z-score at the voxel level: 4.32) and subcortical white matter in the right temporal lobe (MNI coordinate: 30, −42, 12; Z-score at the voxel level: 3.65).
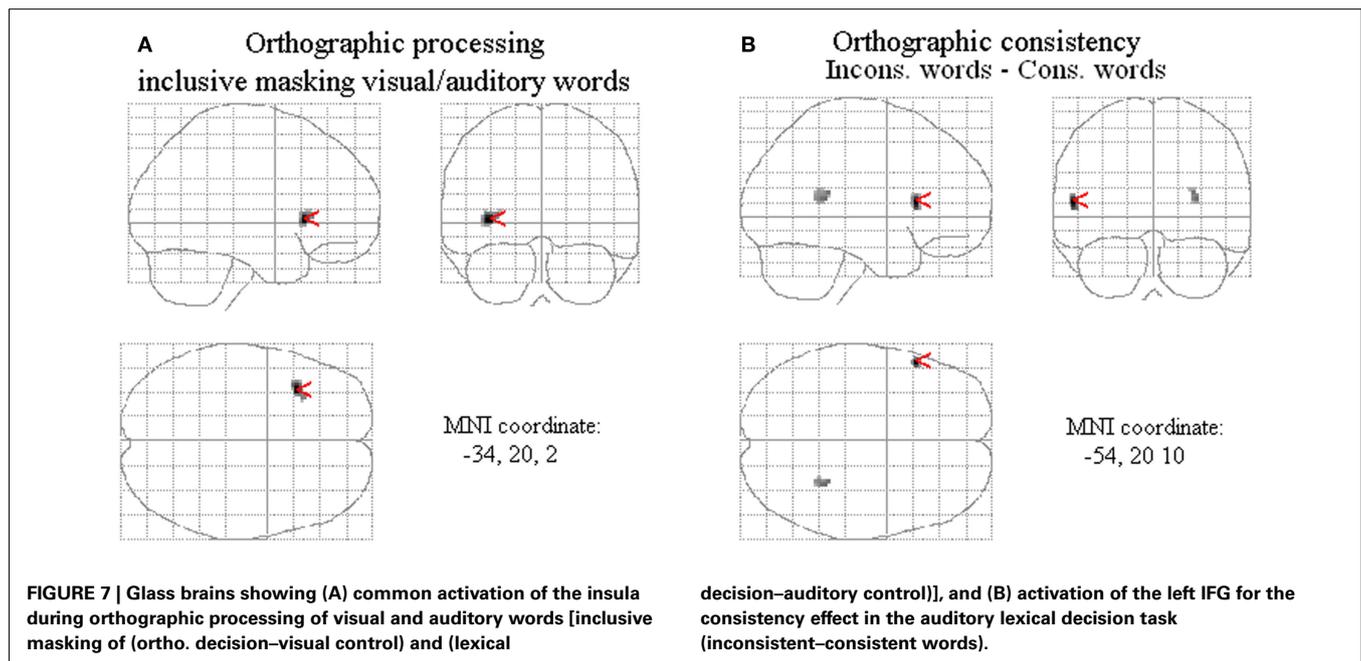
DISCUSSION

In the present study, we combined two fMRI experiments, one in the visual modality whose purpose was to identify the brain areas involved in orthographic processing, and one in the auditory modality that looked for the cerebral bases of the orthographic consistency effect, a behavioral effect that reveals the influence of orthographic information in speech perception. We assumed that the consistency effect either resulted from on-line activation of posterior brain areas or from structural changes in the perisylvian speech network. Finding posterior brain activation in the auditory word recognition task with inconsistent and consistent words was

taken to support the recycling hypothesis whereas activation of the perisylvian region would be favor the phonological restructuring hypothesis.

In the visual modality, we presented pairs of words and yoked pseudohomophones and asked the participants to decide which letter-string was a real word. Given that words and pseudohomophones presented on each trial shared the same phonology, participants had to use lexical orthographic knowledge to make their decision. By contrasting this task to a visual control task, we obtained bilateral activation of the visual cortex in the region of the calcarine sulcus but no activation of the left fusiform gyrus. At face value, the absence of activation in the VWFS would speak against the on-line co-activation account and the recycling hypothesis presented above.

One might be surprised that the VWFA was not activated in the present study since this structure is supposed to be in charge of orthographic processing of visually presented words (for a review,

FIGURE 7 | Glass brains showing (A) common activation of the insula during orthographic processing of visual and auditory words [inclusive masking of (ortho. decision–visual control) and (lexical decision–auditory control)], and **(B)** activation of the left IFG for the consistency effect in the auditory lexical decision task (inconsistent–consistent words).

see Dehaene and Cohen, 2011). This result could be due to the fact that we used the same pairs of items in the orthographic decision and the control tasks. In contrast, most empirical data supporting the idea that the VWFA is in charge of visual word recognition have been obtained by contrasting word-like letter strings (e.g., consonant strings, pseudowords, words) to rest or to simple visual features. In a recent review article, Dehaene and Cohen (2011) insist on the visual nature of the orthographic processes that take place in the VWFA. For example, this region is particularly sensitive to line junctions of letters (Szwed et al., 2011). Dehaene and Cohen also acknowledge that activation of the VWFA depends heavily on the task demands and experimental conditions. They assert that, "to test models of neural coding in the VWFA, it is essential to use short presentation times and minimal tasks that emphasize bottom-up processing (e.g., passive viewing or simple target detection)" (Dehaene and Cohen, 2011; Box 1, p.256). This statement was made in reaction to a challenging paper by Price and Devlin (2011) who claimed that the ventral occipito-temporal cortex (vOTC), that hosts the VWFA, is very sensitive to top-down information and turns out to be "specific" to either words or objects, depending on the task and the nature of the processes in play in the associative cortices.

Our results are neither compatible with Dehaene and Cohen's idea of a specific role of the VWFA in orthographic processing nor with Price and Devlin's top-down view because if activation of the VWFA were top-down driven, we should have observed a BOLD signal difference between the orthographic decision and the control tasks since the orthographic decision task is more orthography-oriented than the visual control task. We believe instead that the present orthographic decision task tapped higher-level (i.e., lexical) orthographic processes that are not really visual in nature. In favor of this position, we obtained a strong activation difference in the left dorsal–anterior insula, which is part of

the perisylvian speech network, beneath Broca's area. This finding would be compatible with the phonological restructuring view.

One might argue that finding insular cortex activation is not specific to word recognition since this structure is known to be activated in multiple linguistic and non-linguistic tasks (e.g., Mutschler et al., 2009). Indeed, the recruitment of the left insula could be due to phonological or decisional processes. However, we believe that a phonological or decisional explanation of this result is not tenable because, in the present study, words and pseudohomophones presented on each trial shared the same phonology and participants had to choose one of these two items in the orthographic decision and the control tasks (i.e., same decisional process). We do believe instead that the activation of the left dorsal–anterior insula is related to higher-level (lexical) orthographic processes required in the orthographic decision task. This interpretation is also supported by the results obtained in the auditory modality (see below).

In the auditory modality, participants were asked to perform lexical decisions on pairs of words and pseudowords. Words were either orthographically inconsistent or consistent. This resulted in the well-established behavioral consistency effect, that is, faster correct responses for consistent words compared to inconsistent words (Ziegler and Ferrand, 1998). At the functional level, we obtained a large activation of the left pre-frontal region, including the IFG and the anterior insula, the left superior and middle temporal gyri, and, to a smaller extent, the left inferior parietal lobule. By comparing the BOLD response to inconsistent and consistent words, we obtained more activation for inconsistent words in the left IFG. Since consistent and inconsistent words only differed on the number of ways their rhyme can be possibly spelled, this finding cannot be attributed to the task (same task for inconsistent and consistent words, same selection mechanisms, same decisional processes), nor to some linguistic variables (see

Materials and Methods). Comparing the areas jointly activated in the visual and auditory word recognition tasks (inclusive masking analysis), we obtained an activation of the left dorsal–anterior insula. Note that the clusters activated in the left IFG and insula (**Figure 7**) are at the same position on the y and z axes, and differ only in terms of depth. As in the orthographic decision task, we found no activation of the VWFA in the auditory modality.

In summary, the experiments that we conducted in the visual and auditory modalities which aimed at identifying the brain regions involved in the processing of orthographic information both point to anterior sites belonging to the perisylvian region: the left dorsal–anterior insula and the left IFG.

At that point, two caveats need to be addressed. First, if orthography is embedded within the spoken language system, how is it possible to find patients for whom orthographic processes, as measured by the ability to make accurate lexical decisions, is spared while phonological and/or semantic processes are deficient (e.g., Lambon Ralph et al., 1998; Blazely et al., 2005)? We believe that such dissociations are possible even if lexical orthography (a word's spelling) were embedded within the spoken language system. The argument is that lexical decisions can be based on low-level orthographic operations that are sensitive to orthographic familiarity. Compared to pseudowords, words have orthographic patterns that occur more frequently, and such orthographic redundancy/familiarity statistics can be used to make accurate lexical decisions in the absence of spoken language (for a similar proposal, see Rogers et al., 2004). As a matter of fact, a wealth of research indicates that the vOTC is the brain region that is sensitive to the orthographic familiarity of letter strings (Binder et al., 2006; Dehaene and Cohen, 2011). However, the extent to which the vOTC actually processes lexical orthography – i.e., whether it hosts the orthographic lexicon – is still a matter of debate (Price and Devlin, 2011). As argued above, the present data suggest that lexical orthographic processes might be "closer" to the spoken language areas than initially thought.

Second, one could argue that the orthographic choice task (decide whether BRANE or BRAIN is a real word) is not as pure of an orthographic measure as one might think because not only orthographic but also phonological activation (at the lexical level) is higher for real words (BRAIN) than for pseudo-homophones (BRANE), as demonstrated by Rastle and Brysbaert (2006). This is an important point, which could explain why we see Broca activation in the orthographic choice task. However, even if there were differences in terms of lexical phonological activation between BRAIN and BRANE, it must be the case that the differences between BRAIN and BRANE are even bigger at the orthographic level (i.e., one is a real word orthographically while the other is not). Thus, while the phonological activation

account could explain why we do see activation in Broca's area, one would still need to explain why we do not see even bigger differences in brain regions that are thought to process orthographic information (e.g., vOTC) given that the orthographic contrast in BRANE–BRAIN pairs is indisputably stronger than their phonological contrast. Thus, the most parsimonious interpretation remains one according to which information about a word's spelling is at least partially processed in Broca's area. Converging evidence for this claim comes from our secondary task, the auditory lexical decision task, in which Broca's area was the only region that differently responded to a purely orthographic manipulation (consistency manipulation).

To conclude, our results support the restructuring hypothesis according to which the speech network is modified in the process of learning to read and code for words' orthography (Perre et al., 2009; Pattamadilok et al., 2010). According to this view, the VWFA or the vOTC would only constitute the visual entry gate to the spoken language system (providing information about letters and orthographically legal sequences of letters) but would not store lexical orthographic knowledge *per se*. This claim is also consistent with neuropsychological data from pure alexic and alexic-plus-agraphic patients. Pure alexia typically results from brain damage to the left vOT and patients are unable to read but can still write. In contrast, alexia-plus-agraphia typically results from a lesion of the speech network, in particular the left angular gyrus, which causes a loss of reading and writing skills. If orthography were exclusively processed in the VWFA, it would be difficult to see why a lesion in the angular gyrus would result in reading loss. Similarly, it would be difficult to explain why a lesion in the VWFA does not preclude writing and spelling aloud.

We do not wish to argue that orthographic knowledge is stored in the insula or in the IFG but rather that orthographic knowledge is distributed over the speech network and is a very part of this network. In our experiments, we observed greater activation of the left insula and IFG probably because these heteromodal regions are hubs of the language network and receive convergent information from many unimodal and heteromodal regions, through long distance connections (e.g., Achard et al., 2006; He et al., 2007). They might integrate the orthographic information coming from unimodal regions involved in the processing of visual or sensori-motor aspects of words, as well as regions involved in the mapping between orthography and phonology. While more research is needed to better understand the intricate relationship between orthography and spoken language processing, the present study suggests that orthographic processing is not restricted to the VWFS but can take place in brain regions, such as Broca, that were previously thought to be dedicated to spoken language processing only.

REFERENCES

Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26, 63–72.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., and Buchanan,

L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33, 739–748.

Blazely, A. M., Coltheart, M., and Casey, B. J. (2005). Semantic impairment with and without surface dyslexia: implications for models of reading. *Cogn. Neuropsychol.* 22, 695–717.

Carr, T. H. (1986). "Perceiving visual language," in *Handbook of Perception and Human Performance*, eds K. R. Boff, L. Kaufman, and J. P. Thomas (New York: Wiley), 29.1–29.82.

Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., and Price, C. J. (2009). An anatomical

signature for literacy. *Nature* 461, 983–986.

Castro-Caldas, A., Petersson, K. M., Reis, A., Stone-Elander, S., and Ingvar, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain* 121(Pt 6), 1053–1063.

Curtin, S., and Werker, J. F. (2007). "The perceptual foundations of phonological development," in *The Oxford Handbook of Psycholinguistics,* ed. G. Gaskell (Oxford: Oxford University Press), 579–599.

Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398.

Dehaene, S., and Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends Cogn. Sci. (Regul. Ed.)* 15, 254–262.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., and Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364.

Ehri, L. C. (1992). "Reconceptualizing the development of sight word reading and its relationship to recoding," in *Reading Acquisition,* eds P. B. Gough, L. E. Ehri, and R. Treiman, (Hillsdale, NJ: LEA), 105–143.

Goswami, U., and Ziegler, J. C. (2006). A developmental perspective on the neural code for written words. *Trends Cogn. Sci. (Regul. Ed.)* 10, 142–143.

Goswami, U., Ziegler, J. C., and Richardson, U. (2005). The effects of spelling consistency on phonological awareness: a comparison of English and German. *J. Exp. Child. Psychol.* 92, 345–365.

He, Y., Chen, Z. J., and Evans, A. C. (2007). Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. *Cereb. Cortex* 17, 2407–2419.

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

Lambon Ralph, M. A., Ellis, A. W., and Sage, K. (1998). Word meaning blindness revisited. *Cogn. Neuropsychol.* 15, 389–400.

McCandliss, B. D., Cohen, L., and Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci. (Regul. Ed.)* 7, 293–299.

Mutschler, I., Wieckhorst, B., Kowalevski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., and Ball, T. (2009). Functional organization of the human anterior insular cortex. *Neurosci. Lett.* 457, 66–70.

Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J., and Devlin, J. T. (2010). How does learning to read affect speech perception? *J. Neurosci.* 30, 8435–8444.

Pattamadilok, C., Morais, J., Ventura, P., and Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: further evidence from French. *Lang. Cogn. Process.* 22, 1–27.

Pattamadilok, C., Perre, L., Dufau, S., and Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *J. Cogn. Neurosci.* 21, 169–179.

Peereman, R., Dufour, S., and Burt, J. S. (2009). Orthographic influences in spoken word recognition: the consistency effect in semantic and gender categorization tasks. *Psychon. Bull. Rev.* 16, 363–368.

Perre, L., Pattamadilok, C., Montant, M., and Ziegler, J. C. (2009). Orthographic effects in spoken language: on-line activation or phonological restructuring? *Brain Res.* 1275, 73–80.

Petersson, K. M., Reis, A., Askelof, S., Castro-Caldas, A., and Ingvar, M. (2000). Language processing modulated by literacy: a network analysis of verbal repetition in literate and illiterate subjects. *J. Cogn. Neurosci.* 12, 364–382.

Price, C. J., and Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci. (Regul. Ed.)* 15, 246–253.

Pulvermuller, F. (1999). Words in the brain's language. *Behav. Brain Sci.* 22, 253–279; discussion 280–336.

Pulvermuller, F., Shtyrov, Y., and Hauk, O. (2009). Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain Lang.* 110, 81–94.

Rastle, K., and Brysbaert, M. (2006). Masked phonological priming effects in English: are they real? Do they matter? *Cogn. Psychol.* 53, 97–145.

Rogers, T. T., Lambon Ralph, M. A., Hodges, J. R., and Patterson, K. (2004). Natural selection: the impact of semantic impairment on lexical and object decision. *Cogn. Neuropsychol.* 21(2/3/4), 331–352.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabregue, R., Amadon, A., and Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage* 56, 330–344.

Taft, M., Castles, A., Davis, C., Lazendic, G., and Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *J. Mem. Lang.* 58, 366–379.

Ventura, P., Morais, J., and Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: from sublexical to lexical involvement. *Cognition* 105, 547–576.

Ventura, P., Morais, J., Pattamadilok, C., and Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Lang. Cogn. Process.* 19, 57–95.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., and Cohen, L. (2007). Dissecting the inner organization of the visual word-form system. *Neuron* 55, 143–156.

Ziegler, J. C., and Ferrand, L. (1998). Orthography shapes the perception of speech: the consistency effect in auditory word recognition. *Psychon. Bull. Rev.* 5, 683–689.

Ziegler, J. C., Ferrand, L., and Montant, M. (2004). Visual phonology: the effects of orthographic consistency on different auditory word recognition tasks. *Mem. Cognit.* 32, 732–741.

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29.

Ziegler, J. C., and Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: a developmental study. *Psychon. Bull. Rev.* 14, 75–80.

Ziegler, J. C., Muneaux, M., and Grainger, J. (2003). Neighborhood effects in auditory word recognition: phonological competition and orthographic facilitation. *J. Mem. Lang.* 48, 779–793.

Ziegler, J. C., Petrova, A., and Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: where do we stand after more than a decade? *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 643–661.