Letters Persistence After Physical Offset: Visual Word Form Area and Left Planum Temporale. An fMRI Study

Francesco Barban,^{1,2*} Gian Daniele Zannino,¹ Emiliano Macaluso,²
Carlo Caltagirone,^{1,3} and Giovanni A. Carlesimo^{1,3}

¹*Clinical and Behavioural Neurology Laboratory, IRCCS Fondazione Santa Lucia, Rome, Italy*

²*Neuroimaging Laboratory, IRCCS Fondazione Santa Lucia, Rome, Italy*

³*Institute of Neurology, University of Rome "Tor Vergata", Italy*

Abstract: Iconic memory is a high-capacity low-duration visual memory store that allows the persistence of a visual stimulus after its offset. The categorical nature of this store has been extensively debated. This study provides functional magnetic resonance imaging evidence for brain regions underlying the persistence of postcategorical representations of visual stimuli. In a partial report paradigm, subjects matched a cued row of a 3×3 array of letters (postcategorical stimuli) or false fonts (precategorical stimuli) with a subsequent triplet of stimuli. The cued row was indicated by two visual flankers presented at the onset (physical stimulus readout) or after the offset of the array (iconic memory readout). The left planum temporale showed a greater modulation of the source of readout (iconic memory vs. physical stimulus) when letters were presented compared to false fonts. This is a multimodal brain region responsible for matching incoming acoustic and visual patterns with acoustic pattern templates. These findings suggest that letters persist after their physical offset in an abstract postcategorical representation. A targeted region of interest analysis revealed a similar pattern of activation in the Visual Word Form Area. These results suggest that multiple higher-order visual areas mediate iconic memory for postcategorical stimuli. *Hum Brain Mapp* 34:1282–1292, 2013. © 2012 Wiley Periodicals, Inc.

Key words: fMRI; iconic memory; partial report; visual perception; consciousness

INTRODUCTION

The persistence of a visual stimulus after its offset has been imputed to a high-capacity low-duration visual memory called iconic memory [Neisser, 1967]. This visual buffer stores rather unlimited information unconsciously

for around half of a second. Differently, visual short-term memory stores limited information consciously for seconds up to minutes [see Offen et al., 2009]. Some authors [see Sligte et al., 2008, 2009; Vandembroucke et al., 2011] argue the existence of an additional visual short-term memory between the two. The nature of the information maintained and the brain substrate underlying this store has been long debated [see for review Coltheart, 1980, 1983; but see also Duncan, 1983]. To our knowledge, only a few neuroimaging studies have been conducted so far to explore the neuronal mechanisms involved in iconic memory [see Ruff et al., 2007; Saneyoshi et al., 2011]. Ruff et al. [2007] showed that the lateral occipital cortex, an area devoted to the processing of the physical stimulus, also underlies its persistence in iconic memory. Further, they

*Correspondence to: Francesco Barban, Neuroimaging Laboratory, IRCCS Fondazione Santa Lucia, Via Ardeatina, 306, 00179 Roma, Italy. E-mail: f.barban@hsantalucia.it

Received for publication 12 January 2011; Revised 13 September 2011; Accepted 11 October 2011

DOI: 10.1002/hbm.21509

Published online 16 January 2012 in Wiley Online Library (wileyonlinelibrary.com).

showed that the right middle frontal gyrus was the only brain area involved in the readout from iconic memory compared to the readout from the physical stimulus. Ruff et al. [2007] presented low-level visual stimuli that can persist only in a precategorical iconic memory. By contrast, iconic memory for categorical stimuli seems to maintain both precategorical and postcategorical representations of visual stimuli [Coltheart, 1980; Duncan, 1983; Townsend, 1973].

Coltheart [1980] argued that the “identity” of a stimulus is more stable than its “physical attributes” and he introduced a distinction between at least two ways in which a stimulus persists after its physical duration. The “visible persistence” maintains a precategorical representation of the stimulus after its offset. This is inversely related to stimulus duration and energy. The “information persistence” maintains an abstract and postcategorical representation of the stimulus after its offset. The duration of the information persistence is longer compared to the duration of the visual persistence and it is unaffected by the stimulus duration and energy. The novelty of the present study is to investigate with functional magnetic resonance imaging (fMRI) the neural basis of iconic memory for stimuli that can be categorized such as letters.

Letters are overlearned visual stylized symbols that access a visual structural description [Hillis and Caramazza, 1995]. This is independent from spatial location, rotation and case. Cohen et al. [2000, 2002] and Dehaene et al. [2002] highlighted a brain area in the ventral occipito-temporal region that is not sensitive to spatial location and to the case of letter strings [Dehaene et al., 2001; 2004]. They referred this region as the visual word form area (VWFA). We predicted that iconic memory for letters would persist within VWFA maintaining the postcategorical information of letters. Within the lateral visual cortex, only precategorical information of stimuli would persist [see Ruff et al., 2007, who presented Landolt rings rather than letters].

In our experiment, we used a modified version of a partial report task [Averbach and Coriell, 1961; Ruff et al., 2007; Sperling, 1960]. This is a kind of paradigm very similar to the one typically used in selective attention tasks. In fact, in our study a visual cue directed exogenous visuospatial attention to one row of a flashed array comprising three competing rows of three elements each. The amount of information contained in the array (either the physical stimulus or its subsequent high capacity short-term iconic memory) exceeded the limit of visual short-term memory [i.e., about four elements, Luck and Vogel, 1997; Sperling, 1960]. Only the cued row (either physical or its iconic memory) was registered in this more durable memory store [Luck and Vogel, 1997]. Once in visual short-term memory, subjects were instructed to compare the cued row with an after coming probe row that appeared at the end of each trial.

To study iconic memory, partial report paradigms manipulate the delay between the array presentation and the cue. Subjects are able to retrieve the same amount of information whether the cue appears simultaneously—or

before—the array (prestimulus cue condition) or whether the cue appears until 250 ms after the array offset (poststimulus cue condition) [see Sperling, 1960]. For this reason, the amount of information contained in iconic memory still exceeds the limit of short-term memory. The prestimulus and poststimulus cue conditions share the same visuospatial attention process [see Ruff et al., 2006]. Critically, visuospatial attention selects a part of the physical stimulus in the prestimulus cue condition. Whereas, visuospatial attention selects a part of the iconic memory of the physical stimulus in the poststimulus cue condition.

In this experiment, we presented a flashed array of letters (postcategorical) or false fonts (precategorical). The task was to recall the cued row of the array. The cue appeared 200 ms after the array was turned off (iconic condition) or when the array appeared (noniconic condition). During the noniconic condition visuospatial attention selected a portion of information from the physical stimulus. Critically, during the iconic condition visuospatial attention selected a portion of information from the iconic memory of the stimulus rather than the visual short-term memory. In fact, not all the information of the array could be registered in short-term memory because the cue delay was too short (200 ms) and the amount of stimuli too high. Only after exogenous visuospatial attention selected the cued information of the array, this was maintained in a more durable memory store such as short-term memory. The main aim of this study was to evidence brain regions sensitive to the interaction between the readout condition (i.e., from iconic memory vs. the physical stimulus) and the kind of visual material presented (letters vs. false fonts: i.e., post/precategorical stimuli).

According to previous studies [Ruff et al., 2007; Saneyoshi et al., 2011], we predicted that the readout from iconic memory, independently from the material involved (letters or false fonts), should recruit the parieto-frontal network involved in selective visuospatial attention. We predicted that the readout from iconic memory would elicit an attentional top-down effect on different brain substrates depending on the precategorical versus postcategorical material. According to the “neurophysiological” perspective [Keysers et al., 2005], the persistence of the brain activity after the stimulus offset may be a general feature of the cortex. Only stimuli that can be categorized such as letters may access to higher visual areas (such as VWFA) that responds to objects familiarity. We believe that the readout from iconic memory of letters and false fonts could show a similar effect for the iconic memory of their noncategorical basic features. These probably persist in V1 or other early visual areas [Ruff et al., 2007]. We predicted that only for postcategorical stimuli as letters, the iconic condition readout would modulate brain areas representing categorical information. A possible candidate for this is the VWFA that processes the abstract representation of letters [see Cohen and Dehaene, 2004]. For this reason, we included in the present study a functional localizer task for the VWFA [see Cohen et al., 2000, 2002].

Figure 1.

Stimuli and behavioral results. **(A)** Schematic representation showing the time course of an iconic and a noniconic memory trial spaced out by an intertrial interval (ITI) of 1,980 ms. In the iconic memory trial, a 3×3 array of letters or false fonts appeared and after 200 ms a cue (two flankers became white) indicated, which of the three rows the subject had to recall from iconic memory. Then a triplet of stimuli appeared below the flankers and the subject had to match it with the cued row of the array. In the noniconic condition, the cue appeared simultane-

ously with the array and the subject had to recall the cued row from the stimulus and to retain it in short-term memory for 300 ms, until the triplet for the matching appeared. **(B)** Percentage of correct responses and mean reaction times for each of the four conditions, iconic memory for letters, noniconic memory for letters, iconic memory for false fonts, noniconic memory for false fonts. Error bars indicate 95% confidence intervals. * $P < 0.05$, *** $P < 0.001$.

MATERIALS AND METHODS

Participants

Fourteen, right-handed [Edinburgh Inventory; Oldfield, 1971], volunteers (nine males and five females; mean age 24.6, range 18–34) participated in the study. None of the subjects had a history of psychiatric or neurological disease or were taking vasoactive or psychotropic medication. All gave their written informed consent. The study protocol was approved by the independent Ethics Committee of the Fondazione Santa Lucia (Scientific Institute for Research, Hospitalisation and Health Care).

Procedure

Stimulus presentation was controlled by a script using Cogent software [Cogent 2000, Functional Imaging Laboratory, Wellcome Department of Imaging Neuroscience, UCL, London] in the MATLAB environment. Participants lay in the scanner in a dimly lit environment and viewed stimuli via a mirror system.

Experimental Task

The aim of this study was to investigate the neural substrates of readout from iconic memory for stimuli that can

be categorized such as letters using a modified version of a standard partial report paradigm [Ruff et al., 2007; Sperling, 1960]. For this purpose, we manipulated two experimental factors orthogonally: experimental condition and material. The factor condition comprised the physical stimulus readout condition (noniconic) and the iconic memory readout condition (iconic). The factor material comprised the presentation of false fonts (precategorical) and letters (postcategorical). Each experimental session lasted about 8 min. In each trial (Fig. 1A), a 3×3 array of visual stimuli (letters or false fonts) was flashed for 100 ms flanked laterally by two columns of three circles; each one was aligned with one of the three rows of the array. The flankers were always present throughout the session. In the iconic condition, 200 ms after the array off go, a horizontal pair of flankers turned white (visual cue) for 100 ms, cueing one of the three rows. In the noniconic condition, the visual cue was presented simultaneously with the array. We decided to present the cue simultaneously rather than before the onset of the array to avoid extra memory for cue location in this condition compared to the iconic condition. The visual cues indicated, which of the three rows of the array (top, middle, or bottom) the subject was attempting to readout and remember. After the cue offset in the iconic condition or after a 300 ms delay from the array + cue offset in the noniconic condition, a triplet of stimuli appeared at the bottom of the screen for 1 sec. The

subject was instructed to match the cued row of the array with the triplet by pressing two different keys with the right hand, one with the index finger if the cued row and the triplet were the same and with the middle finger if they were different.

A modified version of the partial report task was used for several reasons. We used a matching task first to avoid oral responses, which might cause the subjects to make head movements in the scanner, and second because false fonts were not possible to be orally reported. Moreover, we used a salient visual cue to allow a faster allocation of visuospatial attention. Auditory cues or not salient visual cues are decoded through a time consuming procedure because the competition among rows presented simultaneously is biased by voluntary endogenous visuospatial attention (i.e., a top down way). In order to speed up the process of visuospatial selection, we used flashing visual flankers as cue. These are salient visual stimuli that receive attentional priority independently of the intention of the observer (i.e., a bottom-up way). Hence, exogenous attention is immediately allocated to the row of the array surrounded by the flankers [see Theeuwes and Belopolsky, 2010]. Besides, an auditory cue could determine a possible confound interpreting the activity within the superior temporal gyrus. However, the major limit of the matching task was that we were unable to estimate the precise amount of information recalled by subjects.

Arrays consisted of 3×3 stimuli, which subtended 2° and were randomly generated by recombining the first nine consonants of the Italian alphabet, B,C,D,F,G,H,L,M,N, (letters arrays) or nine false fonts created by moving parts of the same nine consonants (false fonts arrays), e.g., . Stimuli were white against a dark background. Each of the four event-type trials resulting from combination of the two main factors (iconic/non-iconic condition; letters/false fonts) was presented 30 times in each experimental session in unpredictable sequences. Same/different triplets and the cued row were equated across trials. If the triplet in the matching task was different from the cued one it was generated by sorting a stimulus (letter or false font) from each of the three rows of the previously presented array and randomly placing it horizontally. We used an event-related design for the experimental task so the different trials resulting from the combination of our conditions were randomly presented. Subjects were subjected to five sessions of 144 trials each in different fMRI runs. Each trial lasted for 1,400 ms and was followed by an intertrial interval (ITI) of 1,980 ms. In addition to the 120 events of interest (30 events per each of the four conditions), each functional run also contained 24 additional null event trials that were randomized with the experimental trials. Before the scanner session, each subject participated in a practice session consisting of 144 trials, which was administered outside the scanner in a psychophysics room.

VWFA Localizer

The VWFA was localized by manipulating two factors orthogonally, that is, material (words and checkerboards) and side of presentation (left or right hemifield) [Cohen et al., 2000, 2002]. The localizer session lasted about 5 min. Subjects were instructed to fixate a permanent central fixation point consisting of a white cross ($1^\circ \times 1^\circ$) against a dark background while stimuli were flashed on the left or the right side. The center of the stimuli from fixation was 3.5° and the maximum eccentricity was 5° . Words used in the localizer consisted in 80 frequent and imageable nouns, which were 2-syllable, 4- to 6-letter words selected from the database developed by Barca et al. [2002]. Words were presented in lowercase and each one appeared once in the left and once in the right hemifield. Checkerboards were small rectangles formed by 16×4 black and white squares. The vertical size of words and checkerboards was $\sim 1^\circ$. Four different types of blocks were presented depending on stimulus type, that is, words or checkerboards, and side of presentation, left or right hemifield. A fifth type of block consisted of the presentation of the fixation point alone in a rest condition. Each block consisted of 20 trials and each trial was 200 ms long with 580 ms interstimulus interval (ISI). The localizer task consisted of 20 blocks performed in a single fMRI run, that is, five condition blocks (including a rest block) repeated four times in pseudorandom order to avoid subsequent repetition of the same kind of block.

fMRI Acquisition

Images were acquired with a T2*-weighted gradient-echo, echo-planar imaging (EPI) sequence on a 3T Siemens Allegra scanner (Siemens Medical Systems, Erlangen, Germany). A quadrature volume head coil was used for radio-frequency transmission and reception. Head movements were restricted by mild restraint and cushioning. Thirty-two axial slices aligned with the bicommissural plane of the functional MR images were acquired in each volume using blood-oxygenation-level-dependent (BOLD) imaging (repetition time = 2.08 s, time echo = 30 ms, flip angle = 70° , matrix size $64 \times 64 \times 32$ voxels, in-plane resolution = 3×3 mm, slice thickness = 2.5 mm, interslice distance = 1.25 mm), covering the entire cortex, including the ventral temporal cortex.

fMRI Analysis

Analysis was performed in SPM5 (Wellcome Department of Cognitive Neurology) as implemented in MATLAB 7.1 (The MathWorks, Natick, MA) for data preprocessing and statistical analyses. For each participant, we acquired 1,385 functional volumes: 160 for the localizer task and 245 for each of the 5 sessions of the experimental task. The first four volumes of each run were discarded to minimize saturation effects and all remaining images were

motion-corrected, with the first volume serving as reference. Slice-acquisition delays were corrected using the middle slice as reference (the 16th slice). All images were normalized to the standard SPM5 EPI template, resampled to 2 mm isotropic voxel size, and spatially smoothed using an isotropic Gaussian Kernel of 8 mm full-width half-maximum.

Statistical inference was based on a random effects approach [Penny and Holmes, 2004]. This involved two levels of analysis. First level. Main experiment: for each subject, the data were best fitted at every voxel using a combination of effects of interest. These were the timing of the four event types in each fMRI-run of the four conditions given by the crossing of our 2×2 factorial design (iconic [letters, false fonts]; noniconic [letters, false fonts]) convolved with the SPM5 hemodynamic response function. Onset of the hemodynamic response function was aligned with onset of the target trial, with a duration = 0. A high-pass filter with a cut-off period of 128 s was used. Linear contrasts were used to determine responses for the four conditions of interest across the five fMRI runs. This resulted in four contrast images per subject. Localizer task: also in this case, for each subject the data were best fitted at every voxel using a combination of effects of interest, which were the timing of the four event types in the fMRI-run of the four conditions given by the crossing of our 2×2 factorial design (words [left, right]; checkerboards [left, right]). The onset of the hemodynamic response function was aligned with the onset of the target, with duration = 0. Linear contrasts were used to determine responses for the four conditions of interest by averaging across the five fMRI runs. This resulted in four contrast images per subject. Second level of the main experiment and the localizer task. The contrast images underwent a within-subjects analysis of variance (ANOVA; implemented in SPM5) that modeled the effect of the four conditions of interest plus the main effect of subjects. Finally, linear compounds were used to compare the condition effects, using between-subjects variance (rather than between scans). Correction for nonsphericity accounted for any differences in error variance across conditions and any nonindependent error terms for the repeated measures [Friston et al., 2002].

For the main experiment, we first highlighted the entire network of brain areas activated by the readout from iconic memory and the network activated by the letters. We tested for the main effect of iconic versus noniconic memory trials (i.e., [iconic/letters + iconic/false fonts] > [noniconic/letters + noniconic/false fonts]) to identify areas involved in the readout from iconic memory. Then, we tested for the main effect of letter versus false font trials (i.e., [iconic/letters + noniconic/letters] > [iconic/false fonts + noniconic/false fonts]) to identify areas involved in orthographic material processing. For these comparisons, SPMs were thresholded at an FWE-corrected cluster-level $P < 0.05$, with the initial voxel-level threshold set to $P < 0.001$ (uncorrected) considering the whole brain as the

volume of interest. Next, we tested our main prediction by looking for the interaction between the two factors to highlight brain activations during readout from iconic memory for letters. We investigated whether readout from iconic memory modulated letters more than false fonts (i.e., [iconic/letters – noniconic/letters] > [iconic/false fonts – noniconic/false fonts]). SPMs were thresholded at an FWE-corrected cluster-level $P < 0.05$, with the initial voxel-level threshold set to $P < 0.001$ (uncorrected). In addition, we checked that all voxels showing this interaction also presented a simple effect of iconic memory for letters (inclusive masking with: iconic/letters > noniconic/letters; $P < 0.05$ uncorrected), thus ensuring that the interaction was not driven solely by the noniconic/false fonts condition. Furthermore, we investigated the interaction between condition (iconic vs. noniconic) and material (letters vs. false fonts) within the VWFA. We identified this area for each subject. First, we isolated in a II level analysis of the localizer task data, the left occipito-temporal area showing the main effect of words versus checkerboards collapsing the side of presentation (i.e., [words left + words right] > [checkerboards left + checkerboards right]) (at $P < 0.05$ uncorrected). Then, we selected the peak of maximum activity (t value maximum) for the contrast words-checkerboard for each subject within this left ventral occipito-temporal area used as search volume (SVC). Finally, we built a 5 mm radius spherical ROI for each subject centred at each individual peak maximum.

We then extracted the percent of signal change (MarsBar 0.41, “MARseille Boite A Région d’Intérêt” SPM toolbox) for each of the four conditions (iconic letters; iconic false fonts; Noniconic letters; Noniconic false fonts) for each subject averaged across sessions. Finally, we performed a repeated measures analysis of variance ANOVA with two within factors: condition (iconic vs. noniconic) and material (letters vs. false fonts).

RESULTS

Behavioral Measures

Figure 1B shows the mean percentage of correct responses and mean reaction times for each of the four event types. A two-way within-participants ANOVA with factor condition (iconic and noniconic) and material (letters and false fonts) performed on accuracy (correct responses/total responses) revealed significant main effects of condition [$F(1,13) = 71.53$, $P < 0.001$] and material [$F(1,13) = 7.95$, $P = 0.01$] and a significant interaction [$F(1,13) = 9.03$; $P = 0.01$]. Similarly, a two-way within-participants ANOVA with the same factors on RT (2 standard deviation outliers were excluded from the analysis in ~5% of the trials) revealed significant main effects of condition [$F(1,13) = 156.97$, $P < 0.001$] and material [$F(1,13) = 55.32$, $P < 0.001$] and a significant interaction [$F(1,13) = 7.9$, $P = 0.01$]. These results indicate that subjects were more accurate and responded more rapidly when the cue was given

Figure 2.

Location for the main effect of readout from iconic memory, orthographic material, and false fonts. **(A)** Enhanced brain activity during readout from iconic memory trials (iconic > non-iconic). On the left side, the anterior regions displayed on the medial and the frontal surface shows activation of bilateral CS and MFG. On the right side, the posterior activations displayed on the lateral surface shows left occipito-temporal (MOG, STS) and parietal (OPS, SPL) regions and the right ITS. **(B)** Enhanced brain activity during letter trials (letters > false fonts) displayed on the lateral surface, showing left temporo-frontal regions (STG, IFG) and on the posterior surface the bilateral cuneus. **(C)** Enhanced brain activity during false font trials (false fonts > letters) displayed on the posterior surface, showing bilateral lateral-occipital regions and the right inferior parietal lobe and on the ventral surface the bilateral fusiform gyrus (FG), the bilateral middle occipital gyrus (MOG) and right superior occipital gyrus (SOG).

simultaneously with the array (noniconic) than when it was given 200 ms after (iconic) and that were more accurate and responded more rapidly to letters than false fonts (see Table in Fig. 1B). Post hoc analysis revealed that the difference between iconic and noniconic condition was significant for both kind of stimuli and for both measures (P consistently <0.001). An inspection of the data (see Fig. 1B) reveals that the difference was greater for letters than for false fonts. This difference was reliable as confirmed by significant interaction effect.

BOLD Activations

Main effect of readout from iconic memory

The first aim in this study was to isolate brain areas involved in the readout from iconic memory by contrasting iconic and noniconic condition trials regardless of whether the matrix was composed of letters or false fonts.

For this purpose, we were interested in prefrontal activations because previous neuroimaging studies report activation of the right middle frontal gyrus [Ruff et al., 2007].

Figure 2A and Table I illustrate regions that showed enhanced responses in iconic memory versus noniconic memory trials. Increased activation was observed in two bilateral and approximately symmetrical anterior clusters

Figure 3.

Location and effect size for readout from iconic memory for orthographic stimuli. Signal plot shows percent of fMRI signal change ($\pm$ standard error of the mean over participants) for each of the four conditions, iconic memory for letters, iconic memory for false fonts, noniconic memory for letters, noniconic memory for false fonts. **(A)** Enhanced brain activity during readout from iconic memory for orthographic stimuli considering the whole brain as region of interest. Left side: Activation of planum temporale is reported on the left lateral surface. This brain region was the only one that showed a significant interaction between iconic memory and letters (i.e., [iconic/letters–noniconic/letters] > [iconic/false fonts–noniconic/false fonts]) and a simple effect of iconic memory for letters (inclusive masking with: iconic/letters > noniconic/letters; P < 0.05 uncorrected). This was done to ensure that the interaction was not driven solely by the noniconic/false fonts condition. Right side: signal plots of the average activity across subjects within a 5 mm radius spherical ROI centred at the main peak within the planum temporale ($x = -52$, $y = -38$; $z = 14$). **(B)** Enhanced brain activity during readout from iconic memory for orthographic stimuli considering the VVFA as region of interest (ROI). Left side: location of 5 mm radius spherical ROIs corresponding to the individual VVFA of each participant displayed on a lateral view. Right side: signal plots of the average of the activity of individual 5 mm radius spherical ROI corresponding to the VVFA.

TABLE I. Main effects and interaction of the main experiment with anatomical locations, significance of activation, extent, and peak coordinates

Contrast	Hemi	Region	Extent _(voxels)	$P_{\text{(corrected)}}$	Z score	Stereotactic coordinates		
						x	y	z
Iconic > Noniconic								
	Left	CS	832	<0.001	5.97	−8	16	50
	Right	CS			5.01	6	16	48
	Left	MFG	528	<0.001	4.57	−28	4	56
	Right	MFG	577	<0.001	4.85	28	2	48
	Left	MOG	859	<0.001	5.15	−40	−80	8
	Left	STS			4.8	−50	−54	12
	Left	OPS	321	<0.01	4.1	−22	−82	32
	Left	SPL			3.75	−18	−70	52
	Right	ITS	201	<0.05	4.55	48	−72	6
Letters > False fonts								
	Left	STG	1379	<0.001	5.34	−62	−42	18
	Left	IFG	247	=0.015	4.28	−58	14	28
	Left	Cuneus	351	=0.003	4.56	−8	−102	18
	Right	Cuneus			4.27	10	−98	24
Iconic Letters–Noniconic Letters > Iconic False fonts–Noniconic False fonts ^a								
	Left	Planum temporale	195	=0.039	4.1	−60	−24	6
					3.43	−52	−38	14

Note: Statistical threshold was set at $P < 0.05$ corrected for multiple comparisons at cluster level, considering the whole brain as the volume of interest. Statistical threshold for cluster extension was set at $P < 0.001$ uncorrected for multiple comparisons. Abbreviations: CS, cingulate sulcus; MFG, middle frontal gyrus; MOG, middle occipital gyrus; STS, superior temporal sulcus; OPS, occipito parietal sulcus; SPL, superior parietal lobe; ITS, inferior temporal sulcus; STG, superior temporal gyrus; IFG, inferior frontal gyrus. Coordinates of activation peaks are given in the MNI stereotactic space.

^aInclusive masking with: iconic/letters > noniconic/letters; $P < 0.05$ uncorrected.

located in the middle frontal gyrus and the cingulate sulcus. In the posterior region, greater activity for the same contrast was observed in the left hemisphere, specifically in the middle occipital gyrus, the superior temporal sulcus, the occipito-parietal sulcus, and the superior parietal lobe and in the right hemisphere in the inferior temporal sulcus.

Main effect of letters

We identified brain areas that process letters (Fig. 2B and Table I) by directly contrasting trials with letters and with false fonts, irrespective of whether or not the subjects were reading from iconic memory or from the physical stimulus. As letters differ from false fonts at various levels of processing besides the visual one, we expected to find enhanced activity in the left hemisphere brain areas dedicated to language processing at different levels, including the phonological and the articulatory processing of orthographic stimuli. Indeed, increased activation was observed in the left hemisphere in the superior temporal gyrus surrounding the posterior sylvian fissure and the inferior frontal gyrus, as well as in the bilateral posterior cuneus. This comparison did not show any brain areas of increased activation in the ventral occipito-temporal region, not even in the VWFA. False fonts versus letters showed enhanced brain activity in a bilateral occipito-temporal region (Fig. 2C) extending in

the ventral and a lateral portion. The cluster in the ventral left hemisphere was medial respect to the VWFA, whereas the cluster in the ventral right hemisphere comprised the entire ventral stream.

Iconic memory for letters

We investigated our critical question concerning the interaction between the two factors, (iconic vs. noniconic) and material (letters vs. false fonts), by looking for areas in which the readout from iconic memory selectively influenced letters versus false fonts. We hypothesized that a possible candidate may be the VWFA. We tested for regions in which the difference between letter iconic memory trials and letter noniconic memory trials was greater than the difference between false font iconic memory trials and false font noniconic memory trials. At the whole-brain level, this comparison revealed a single cluster active in the left planum temporale (Fig. 3A and Table I) within the left superior temporal gyrus in the depth of the sylvian fissure.

Iconic Memory for Letters in VWFA

Since we predicted a possible involvement of the VWFA in the readout from iconic memory for letters, to individuate this area for each individual subject we performed a localizer task and we extracted the percent of signal

change for each trial type (iconic letters; iconic false fonts; noniconic letters; noniconic false fonts) for each subject averaged across sessions (Fig. 3B). Then, we performed a repeated measures analysis of variance ANOVA with two within factors: condition (iconic vs. noniconic) and material (letters vs. false fonts).

The analysis revealed a nonsignificant main effect of iconic memory versus noniconic memory ($F(1,13) = 0.863$, $P = 0.370$), a nonsignificant effect of letters versus false fonts ($F(1,13) = 0.168$, $P = 0.689$), and a statistical trend for the interaction between these two factors ($F(1,13) = 3.310$, $P = 0.092$). A one-sample T-test on the relevant weighted difference ([iconic/letters – noniconic/letters] – [iconic/false fonts – noniconic/false fonts]) demonstrated that the iconic memory effect was significantly larger in letter than false fonts ($T(13) = 1.817$, $P = 0.046$ one-tailed).

DISCUSSION

Although iconic memory has been extensively investigated [for reviews, see Coltheart, 1983; Long, 1980], only a few neuroimaging studies have been conducted so far about the brain substrate underpinning this high-capacity low-duration visual memory store [see Ruff et al., 2007; Saneyoshi et al., 2011]. Moreover, the nature of the information that persists in this store has been long debated [Coltheart, 1983; Duncan, 1983]. This fMRI study investigated the neural substrates underlying the persistence of visual stimuli that can be categorized, such as letters. The main aim was to find brain areas showing an interaction between the source of the attentional readout (iconic vs. noniconic) $\times$ material (letters vs. false fonts). This corresponded to the brain area where a greater iconic/noniconic readout modulation was induced by the presentation of letters compared to false fonts. According to Coltheart [1980], the abstract information persists longer compared to the visible persistence that fades away more rapidly. Therefore, we expected that the iconic memory readout would exert its top-down influence during the presentation of postcategorical stimuli (letters) on high-order visual areas that process the abstract visual information.

A very intriguing result of this study was to find a significant interaction between condition (iconic/noniconic) and material (letter/false fonts) in the left planum temporale in the whole-brain voxel-wise analysis. This suggested that the left planum temporale holds a representation of letters that persists after the physical stimulus has been removed. This is consistent with Keyers et al. [2005, p 331], who showed that also in high-order temporal areas visual stimuli shortly persist after their offset. About the nature of the representation of letters persisting within the planum temporale we cannot be conclusive. Although the planum temporale has been traditionally considered as an auditory-motor interface, recent investigations have shown that this area is also involved in written language [Buchsbaum et al., 2005; Jobard et al., 2003]. In particular, it has

been defined as a “computational hub” [Griffiths and Warren, 2002] responsible for processing and matching incoming acoustic and visual patterns with acoustic pattern templates [Nakada et al., 2001], integrating spoken and written language [van Atteveldt et al., 2004]. Moreover, this region is related with the learning of new audiovisual associations [Hasegawa et al., 2004] useful for the acquisition of reading skills [Blau et al., 2010]. Therefore, we argue that abstract visual memories of letters persist within the left planum temporale and that they match with acoustic templates of speech sounds.

In the whole-brain voxel-wise analysis, a network of brain areas showed enhanced activity corresponding to the readout from iconic memory (iconic > noniconic condition) collapsing letters and false fonts. This pattern was present in anterior regions of both hemispheres, specifically in the middle frontal gyrus and the cingulate sulcus. Posteriorly, the left lateral occipito-temporal region, the superior parietal lobe and the right lateral temporal lobe also showed the same pattern of activation. These results agree with previous studies [Nobre et al., 2004; Ruff et al., 2007; Saneyoshi et al., 2011] assuming that a fronto-parietal network is responsible to the readout operated by visuo-spatial attention to iconic memory. The activation in the right middle frontal gyrus that we observed in our data was more posterior to the one described in Ruff et al. [2007]. This area, together with superior parietal areas, has been suggested to play a role for the access of visual representations to consciousness [see Block, 2005]. This is in agreement with the unconscious nature of iconic memory [Sperling, 1967]. In fact, during a partial report task, only the cued part of the information contained in iconic memory becomes conscious to the subject. Finally, we found a left sided fronto-temporal network of brain areas showing enhanced activity corresponding to the letters presentation (letters > false fonts) collapsing iconic and noniconic condition. This activity reflects the recruitment of brain regions usually involved during reading [Jobard et al., 2003]. Moreover, the bilateral cuneus showed enhanced brain activity during letters presentation. This is consistent with previous fMRI results [James and Gauthier, 2006] showing that different letter processing tasks activate a network of brain regions including the left cuneus. In particular, this brain region is involved in tasks of perception of letters and also during visual imagery of letters. James and Gauthier [2006] argued that this region may not be letter-selective and its activity may be related more in general to drawing presentation. In this perspective, we cannot be conclusive about the interpretation of our results. Since letters and false fonts are very similar line drawings, our result may be driven by low-level image differences between these two kinds of material used. The main effect of letters versus false fonts did not show enhanced activity within the VWFA. Letters are visually similar to false fonts and for this reason, a main effect contrasting the activity related to the two kinds of stimuli could be inefficient to show domain specificity for letters. Instead, differences

between letters and false fonts could emerge observing an interaction between different tasks and these two kinds of stimuli. In fact, this is what we obtained in the analysis restricted to the VWFA. We functionally defined the VWFA region with separate localizer scans for each subject. This was the left side region on the ventral occipito-temporal area responding more to words than to checkerboards irrespective of the side (left/right) of presentation. This more targeted ROI analysis that specifically considered the (functionally defined) VWFA, showed an almost significant modulation during the readout from iconic memory compared with the readout from the physical stimulus for letters compared with false fonts. These results confirmed our expectations about the VWFA as a brain region underlying the persistence of postcategorical stimuli as letters. The presence of an abstract representation for letters within this area allows an automatic processing of this material when the letters are physically present. This requires a weaker attentional readout effort. The readout effect is in turn enhanced when a more instable iconic memory of the stimulus is present. Precategorical stimuli such as false fonts require the same attentional readout in both conditions, as they have no abstract representation that can facilitate their processing when the stimulus is present (noniconic condition). Moreover, no abstract representation persists when the stimulus has been physically removed (iconic condition). The interaction approached significance only when the analysis was confined to the VWFA as region of interest. This could be probably due to intersubjects variability in the functional-anatomical correspondence of the VWFA [e.g., see Cohen et al., 2000]. In fact, individually defined VFAs of only four subjects overlapped (Fig. 3B) within a region corresponding to the VWFA described by Cohen et al. [2000]. The localization of VFAs of the others subjects varied antero-posteriorly along the occipito-temporal sulcus, that is the anatomical landmark of this functionally defined region [Cohen and Dehaene, 2004]. For this reason, we were quite confident that we did not include in the analysis any retinotopic brain regions that are usually located more medially [see Wandell and Winawer, 2011]. VWFA is far from being a homogeneous and definite structure specialized in a recent cultural domain such as reading [Dehaene and Cohen, 2007]. This could explain the variability of the VWFA localization between subjects. The process of reading acquisition tunes the left occipito-temporal cortex to a gradient of increasingly complex neuronal detectors from individual letters to morphemes [Vinckier et al., 2007] and this process is probably influenced by individual differences. Moreover, participants to this study were Italian readers. This is a language with a consistent orthography where the letter-to-sound conversion predominates over direct matching between a visual form of words and their meaning [Demonet et al., 2005]. This might produce a weaker specialization of the VWFA in these readers.

The ROI analysis within this area revealed no difference between letters and false fonts. This result may be due to

the use in this study of arrays of consonants. In fact, Vinckier et al. [2007] reported that strings of consonants not following graphotactic rules and false fonts produce a similar effect within the VWFA. Moreover, pseudowords, letter and false fonts strings activate a largely overlapping network including also the ventral occipito-temporal region [Tagamets et al., 2000]. For this reason, as we discussed above, a main effect contrasting the activity related to the two kinds of stimuli could be inefficient to show domain specificity for letters. We conclude that brain activity within VWFA was modulated by the source of the readout (iconic memory vs. physical stimulus) only when postcategorical stimuli as letters were presented. Therefore, this area stores abstract memories only for orthographic stimuli [Cohen et al., 2002; Price et al., 1996] that persist after the stimulus has been turned off.

The role of iconic memory has been long debated [Long, 1980]. Our results suggest its involvement in reading. The VWFA and planum temporale are brain areas of the ventral (orthographic) and the dorsal (phonological) systems involved in this cultural ability [Bolger et al., 2005; Jobard et al., 2003]. During text reading, iconic memory could allow the integration of information acquired in subsequent fixations, the so-called trans-saccadic integration [Irwin, 1991; Prime et al., 2007]. Moreover, previous studies have shown that word recognition was not suppressed while the eyes were in motion [Irwin, 1998; Yatabe et al., 2009]. Iconic memory could maintain orthographic material during saccades to enhance reading skills. However, further investigation needs to be carried out in order to address those hypotheses.

In conclusion, our results provide evidence that readout from iconic memory of letters activates an area of the left superior temporal gyrus within the planum temporale. A similar pattern of results emerged also when the analysis was restricted to the VWFA [Cohen et al., 2000] that was individually isolated in our subjects through a dedicated functional localizer. Our results may be interpreted within a “neurophysiological” perspective [Keysers et al., 2005]. This view considers iconic memory as a “multilayered process” with different time courses for the different degrees of categorization of visually presented stimuli. Moreover, some authors suggest [see Vandenbroucke et al., 2011] the existence of different visual short-term memories in between iconic memory and visual working memory with different durations. Further examination is needed to understand whether different structures underlie precategorical and postcategorical stimuli persistence after their offset. These structures might be disposed along a continuum that goes from the early visual brain areas [Ruff et al., 2007] to high-order visual and multimodal areas as showed in this study.

REFERENCES

- Averbach E, Coriell AS (1961): Short-term memory in vision. *Bell Syst Tech J* 40:309–328.

- Barca L, Burani C, Arduino LS (2002): Word naming times and psycholinguistic norms for Italian nouns. *Behav Res Methods Instrum Comput* 34:424–434.
- Blau V, Reithler J, van Atteveldt N, Seitz J, Gerretsen P, Goebel R, Blomert L (2010): Deviant processing of letters and speech sounds as proximate cause of reading failure: A functional magnetic resonance imaging study of dyslexic children. *Brain* 133:868–879.
- Block N (2005): Two neural correlates of consciousness. *Trends Cogn Sci* 9:46–52.
- Bolger DJ, Perfetti CA, Schneider W (2005): Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Hum Brain Mapp* 25:92–104.
- Buchsbaum BR, Olsen RK, Koch PF, Kohn P, Kippenhan JS, Berman KF (2005): Reading, hearing, and the planum temporale. *Neuroimage* 24:444–454.
- Cohen L, Dehaene S (2004): Specialization within the ventral stream: The case for the visual word form area. *Neuroimage* 22:466–476.
- Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, Henaff MA, Michel F (2000): The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123:291–307.
- Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S (2002): Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125:1054–1069.
- Coltheart M (1980): The persistences of vision. *Philos Trans R Soc Lond B Biol Sci* 290:57–69.
- Coltheart M (1983): Iconic memory. *Philos Trans R Soc Lond B Biol Sci* 302:283–294.
- Dehaene S, Cohen L (2007): Cultural recycling of cortical maps. *Neuron* 56:384–398.
- Dehaene S, Naccache L, Cohen L, Bihan DL, Mangin JF, Poline JB, Riviere D (2001): Cerebral mechanisms of word masking and unconscious repetition priming. *Nat Neurosci* 4:752–758.
- 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* 13:321–325.
- Dehaene S, Jobert A, Naccache L, Ciuciu P, Poline JB, Le Bihan D, Cohen L (2004): Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychol Sci* 15:307–313.
- Demonet JF, Thierry G, Cardebat D (2005): Renewal of the neurophysiology of language: Functional neuroimaging. *Physiol Rev* 85:49–95.
- Duncan J (1983): Perceptual selection based on alphanumeric class: Evidence from partial reports. *Percept Psychophys* 33:533–547.
- Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J (2002): Classical and Bayesian inference in neuroimaging: Applications. *Neuroimage* 16:484–512.
- Griffiths TD, Warren JD (2002): The planum temporale as a computational hub. *Trends Neurosci* 25:348–353.
- Hasegawa T, Matsuki K, Ueno T, Maeda Y, Matsue Y, Konishi Y, Sadato N (2004): Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Brain Res Cogn Brain Res* 20:510–518.
- Hillis AE, Caramazza A (1995): Cognitive and neural mechanisms underlying visual and semantic processing: Implications from "Optic Aphasia". *J Cogn Neurosci* 7:457–478.
- Irwin DE (1991): Information integration across saccadic eye movements. *Cogn Psychol* 23:420–456.
- Irwin DE (1998): Lexical processing during saccadic eye movements. *Cogn Psychol* 36:1–27.
- James KH, Gauthier I (2006): Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia* 44:2937–2949.
- Jobard G, Crivello F, Tzourio-Mazoyer N (2003): Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage* 20:693–712.
- Keyser C, Xiao DK, Foeldiák P, Perrett DI (2005): Out of sight but not out of mind: The neurophysiology of iconic memory in the superior temporal sulcus. *Cogn Neuropsych* 22:316–332.
- Long GM (1980): Iconic memory: A review and critique of the study of short-term visual storage. *Psychol Bull* 88:785–820.
- Luck SJ, Vogel EK (1997): The capacity of visual working memory for features and conjunctions. *Nature* 390:279–281.
- Nakada T, Fujii Y, Yoneoka Y, Kwee IL (2001): Planum temporale: Where spoken and written language meet. *Eur Neurol* 46:121–125.
- Neisser U. 1967. *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Nobre AC, Coull JT, Maquet P, Frith CD, Vandenberghe R, Mesulam MM (2004): Orienting attention to locations in perceptual versus mental representations. *J Cogn Neurosci* 16:363–373.
- Offen S, Schluppeck D, Heeger DJ (2009): The role of early visual cortex in visual short-term memory and visual attention. *Vision Res* 49:1352–1362.
- Oldfield RC (1971): The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Penny W, Holmes A (2004): Random effects analysis. In: Frackowiak R, et al., editors. *Human brain function II*. San Diego, CA: Elsevier. p 843–851.
- Price CJ, Moore CJ, Frackowiak RS (1996): The effect of varying stimulus rate and duration on brain activity during reading. *Neuroimage* 3:40–52.
- Prime SL, Tsotsos L, Keith GP, Crawford JD (2007): Visual memory capacity in transsaccadic integration. *Exp Brain Res* 180:609–628.
- Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes JD, Rees G, Josephs O, Deichmann R, Driver J (2006): Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Curr Biol* 16:1479–1488.
- Ruff CC, Kristjansson A, Driver J (2007): Readout from iconic memory and selective spatial attention involve similar neural processes. *Psychol Sci* 18:901–909.
- Saneyoshi A, Niimi R, Suetsugu T, Kaminaga T, Yokosawa K (2011): Iconic memory and parietofrontal network: fMRI study using temporal integration. *Neuroreport* 22:515–519.
- Sligte IG, Scholte HS, Lame VAF (2008): Are there multiple visual short-term memory stores? *Plos One* 3:e1699.
- Sligte IG, Scholte HS, Lame VAF (2009): V4 activity predicts the strength of visual short-term memory representations. *J Neurosci* 29:7432–7438.
- Sperling G (1960): The information available in brief visual presentations. *Psychol Monogr* 74:1–29.
- Sperling G (1967): Successive approximations to a model for short term memory. *Acta Psychol (Amst)* 27:285–292.
- Tagamets MA, Novick JM, Chalmers ML, Friedman RB (2000): A parametric approach to orthographic processing in the brain: An fMRI study. *J Cogn Neurosci* 12:281–297.
- Theeuwes J, Belopolsky A (2010): Top-Down and Bottom-Up Control of Visual Selection: Controversies and Debate. In: Coltheart V, editor. *Tutorials in Visual Cognition*. New York: Psychology Press. p 67–91.

- Townsend VM (1973): Loss of spatial and identity information following a tachistoscopic exposure. *J Exp Psychol* 98:113–118.
- van Atteveldt N, Formisano E, Goebel R, Blomert L (2004): Integration of letters and speech sounds in the human brain. *Neuron* 43:271–282.
- Vandenbroucke ARE, Sligte IG, Lamme VAF (2011): Manipulations of attention dissociate fragile visual short-term memory from visual working memory. *Neuropsychologia* 49: 1559–1568.
- Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L (2007): Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron* 55:143–156.
- Wandell BA, Winawer J (2011): Imaging retinotopic maps in the human brain. *Vision Res* 51:718–737.
- Yatabe K, Pickering MJ, McDonald SA (2009): Lexical processing during saccades in text comprehension. *Psychon Bull Rev* 16:62–66.