Dissociating visual form from lexical frequency using Japanese

Tae Twomey [a,*], Keith J. Kawabata Duncan [b,c], John S. Hogan [a], Kenji Morita [c], Kazumasa Umeda [c], Katsuyuki Sakai [c], Joseph T. Devlin [a]

[a] Cognitive, Perceptual and Brain Sciences, UCL, London, UK
[b] Wellcome Trust Centre for Neuroimaging, UCL, London, UK
[c] Department of Cognitive Neuroscience, University of Tokyo, Tokyo, Japan

ARTICLE INFO

ABSTRACT

In Japanese, the same word can be written in either morphographic Kanji or syllabographic Hiragana and this provides a unique opportunity to disentangle a word's lexical frequency from the frequency of its visual form – an important distinction for understanding the neural information processing in regions engaged by reading. Behaviorally, participants responded more quickly to high than low frequency words and to visually familiar relative to less familiar words, independent of script. Critically, the imaging results showed that visual familiarity, as opposed to lexical frequency, had a strong effect on activation in ventral occipito-temporal cortex. Activation here was also greater for Kanji than Hiragana words and this was not due to their inherent differences in visual complexity. These findings can be understood within a predictive coding framework in which vOT receives bottom-up information encoding complex visual forms and top-down predictions from regions encoding non-visual attributes of the stimulus.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Of all the world's languages, Japanese uniquely relies on multiple written scripts for its everyday use. Books, magazines, and advertisements all mix morphographic[1] Kanji with syllabographic Hiragana such that no adult text consists solely of one script (although some children's books are written only in Hiragana). As a result, Japanese adults are equally familiar with both scripts and individual words can be written multiple ways. For instance, a word such as "apple" is as common in Hiragana (りんご) as Kanji (林檎). In many cases, however, one form will be more common than the other. For instance, "mischief" is usually written in Hiragana (わんぱく) but sometimes occurs as Kanji (腕白). The fact that the same word can be written in different scripts means that Japanese offers a unique opportunity to disentangle the frequency of a written word (i.e. its lexical frequency) from the frequency of its visual form (i.e. its visual familiarity). In alphabetic languages, on the other hand, written lexical frequency and visual familiarity are essentially the

same thing – both measure how frequently a word appears in print. Consequently, much of the literature focuses on lexical frequency as a key factor in understanding the nature of the neural information processing in brain regions related to reading (Fiebach, Friederici, Müller, & Von Cramon, 2002; Hauk, Davis, & Pulvermüller, 2008), despite the possibility that visual familiarity and lexical frequency may have differential effects.

This is particularly relevant to theories of the left ventral occipito-temporal cortex (vOT) – an area consistently engaged by visual word recognition (Price & Mechelli, 2005). Here, activation is greater for low than high frequency words (Chee, Westphal, Goh, Graham, & Song, 2003; Hauk et al., 2008; Joubert et al., 2004; Kronbichler et al., 2004) and this effect is not limited to alphabetic languages but also seen in Chinese (Kuo et al., 2003; Lee et al., 2004). Some accounts claim that the area stores and processes language-specific information (Glezer, Jiang, & Riesenhuber, 2009; Kronbichler et al., 2004). Indeed, Kronbichler and colleagues (2004) have argued that vOT is the site of the "orthographic input lexicon" where orthographic word representations that abstract away from details of the visual form are stored (Bruno, Zumberge, Manis, Lu, & Goldman, 2008; Glezer et al., 2009; Kronbichler et al., 2004, 2007; van der Mark et al., 2009). Like logogens (Morton, 1969), entries in this orthographic input lexicon are sensitive to experience, with access to less frequent words requiring greater effort and therefore resulting in greater activation. An alternate explanation is that vOT represents visual form more generally and is not specialized for written words (Price & Devlin, 2011). By this account, the same vOT neurons that represent spatial

[1] Although Japanese Kanji and Chinese Hanzi characters are visually identical, we use the term "morphographic" when referring to Kanji characters because most characters function as grammatical morphemes. We use the term "logographic" for Chinese characters since this is the term widely used in the literature, even though the characters, too, are largely morphographic (Rogers, 2005). In contrast, Hiragana characters refer to individual mora, a phonological timing unit which largely corresponds to a syllable in Japanese.

configurations important to written words also contribute to other visual stimuli such as objects, scenes and faces. Reciprocal connections with higher order association areas link these visual representations with non-visual properties of the stimulus such as its sound (phonology) or meaning (semantics). As a result, frequency effects arise from the interaction of bottom-up and top-down constraints. Specifically, high frequency written words are more familiar visual patterns and thus have more accurate top-down predictions into vOT reducing prediction error and therefore activation. In contrast, low frequency words result in greater prediction error, increasing the processing demands on vOT and thereby increasing the activation (Price & Devlin, 2011).

The aim of the current study was to test this Interactive Account using Japanese to differentiate between the frequency of a word and its visual familiarity. This distinction is not possible in most languages as individual words can only be correctly written one way. For instance, in alphabetic languages like English, it is possible to write a word phonetically (e.g. "brane") but literate readers immediately recognize these as incorrectly spelled which is very different from seeing two different forms of a correctly spelled word. Another possibility would be to test bilinguals with the same word in different languages (e.g. "米" in Japanese and "rice" in English) but again, this is not optimal since the difference in script is confounded by a difference in language. In this example, the semantic properties of "米" and "rice" are not identical because unlike the English word, the Japanese means "uncooked rice" and there is a separate word for "cooked rice". In other words, the "same" word in two different languages often are subtly different, confounding differences in their visual form. Therefore, only those languages that allow a word to be written in multiple forms suffice for dissociating visual familiarity from lexical frequency. Although Chinese (Hanzi vs. Pinyin) and Korean (Hanja vs. Hangul) offer this possibility, in these languages only one form is in daily use, making one script much more familiar than the other. Japanese, on the other hand, is unique in its reliance on multiple forms for everyday use and therefore provides a fertile ground for testing theories of vOT function in reading.

According to the Interactive Account, visual familiarity is expected to strongly modulate activation in vOT, consistent with the hypothesis the region plays a more domain-general role in representing visual patterns, of which written words are only one example. Lexical frequency, on the other hand, is predicted to interact with visual familiarity such that only low frequency words are affected by visual familiarity. For highly frequent words, vOT will receive sufficiently accurate top-down predictions to quickly and accurately match the bottom-up visual information regardless of whether the visual form is more or less familiar. In contrast, less frequent words will send less accurate top-down predictions to vOT, resulting in greater prediction error. Consequently, activation in vOT for low frequency words will benefit from greater visual familiarity. Finally, the Interactive Account makes one further prediction. Namely, consistent with previous studies (Ha Duy Thuy et al., 2004; Ino, Nakai, Azuma, Kimura, & Fukuyama, 2009; Nakamura, Dehaene, Jobert, Le Bihan, & Kouider, 2005; Sakurai et al., 2000), it predicts that greater vOT activation for Kanji relative to Hiragana, due to differences in the top-down signals. Specifically, the inconsistent mapping between Kanji characters and their phonology results in greater prediction error than Hiragana words where there is a consistent mapping between characters and syllables. Consequently, the magnitude of the BOLD signal will be greater for Kanji than for Hiragana.

The current experimental design aimed to evaluate these hypotheses. Participants performed a lexical decision task where they decided whether visual stimuli represented real Japanese words. Stimuli were written in either Kanji or Hiragana and fully crossed with visual familiarity. In other words, each word appeared twice in the course of the experiment, once in Kanji and once in Hiragana. Half of the words were more commonly written in Kanji while the other half were more common in Hiragana. This design allowed us to look for main effects of Visual Familiarity (high vs. low) and Script (Kanji vs. Hiragana) as well as their potential interaction. A second analysis recoded all the words into four sets divided according to lexical frequency (i.e. independent of script) and looked for frequency effects and their interaction with visual familiarity. In this fashion, we could independently evaluate the effects of visual familiarity and lexical frequency on vOT as well as look for potential script differences.

2. Methods

2.1. Participants

Forty native Japanese speakers participated in this study although the data from six were excluded due to either excessive motion inside the scanner (i.e. motion greater than the dimensions of a voxel; $n = 3$ participants) or due to poor performance (i.e. accuracy less than 60% in one or more conditions; $n = 3$ participants). Consequently, only data from 34 participants (13M, 21F, aged 21–62) were included in the final analyses. Since we hypothesized that the amount of exposure to written Japanese may affect the activation, we tested two groups of native Japanese readers: university students in Tokyo with daily exposure to written Japanese ($n = 15$, 10M, 5F, aged 21–31) and Japanese ex-patriots who had lived outside Japan for a minimum of 3 years and thus had reduced exposure to written Japanese in their daily lives ($n = 19$, 3M, 16F, 29–62). The imaging analyses, however, revealed no significant interactions between Group (Tokyo vs. London) and any other factor. As a result, the results presented here collapse over Group despite including it as a factor in the analyses to better model structured variance in the data.

All participants were native Japanese speakers born and educated in Japan through at least secondary school. Consequently, all were literate adult readers in Japanese familiar with both Kanji and Hiragana. In addition, all were right-handed except for one who was confirmed to be ambidextrous according to the Edinburgh Handedness Inventory (Oldfield, 1971). None reported a history of reading difficulties or neurological problems. Each of the London participants had lived outside of Japan, China and Korea (where morphographic or logographic scripts are used) for at least for 3 years (range: 3–34 years, mean = 11). Testing in Tokyo was approved by the ethics committees of the Graduate School of Medicine, the University of Tokyo (#2968), and the ethics committee of the Brain Science Institute, Tamagawa University (C21–4). In London, ethical approved was granted by the NHS Berkshire Research Ethics Committee (06/Q1602/20).

2.2. Experimental procedures

The participants' task was to view strings of characters and decide whether the string formed an existing Japanese word or not. The task involved 60 words, each of which was presented twice – once in Kanji and once in Hiragana. One half of the words are most commonly written in Kanji and the other half are most commonly written in Hiragana. An equal number of nonwords, divided evenly between Kanji and Hiragana, were included to ensure adequate task performance.

A trial began with a fixation cross presented for 500 ms. A stimulus (written horizontally using the MS Gothic font) was then presented for 500 ms, followed by a jittered inter-stimulus interval of 1–4 s (mean = 2.5 s). Therefore, the average trial length was 3.5 s. Stimuli were presented in a blocks of 15 trials (lasting 54 s) which

included both "yes" and "no" responses in a pseudorandomized order. These were separated by 15 s blocks of fixation that served as an implicit baseline. Over a run, there were eight blocks of task and eight blocks of rest. Therefore, each run lasted 9 min and 12 s. There were two runs. Responses were made with a button press, using either the index or middle finger of their right hand to indicate "yes" and "no." The response fingers were fully counter-balanced across participants. The stimuli were projected onto a screen and viewed via mirrors attached to the head coil. Participants practiced each task inside the scanner before the main runs began. No items that were used in the practice runs occurred during the main experiment.

2.3. Stimuli

The word stimuli were obtained from the NTT Japanese Psycholinguistic Database (Amano & Kondo, 2003a, 2003b) by identifying 30 words that had a higher visual familiarity score when written in Kanji than any other script (including Hiragana, Katakana and mixed scripts). Another set of 30 words was found that had a higher visual familiarity score when written in Hiragana than any other script. The Kanji words were then transliterated into Hiragana ("展望"→" てんぼう") and the Hiragana words transliterated into Kanji ("とんち"→ 頓知") producing 120 words (60 written in Kanji, 60 written in Hiragana). It is important to note that while this transcription changes the visual familiarity of the word, the lexical frequency of the word remains constant. All words were 2 or 3 characters in length when written in Kanji, and between 2 and 6 when written in Hiragana. The resulting word set had four different conditions, each with 30 items (high visual familiarity Kanji words, low visual familiarity Kanji words, high visual familiarity Hiragana words, and low visual familiarity Hiragana words) corresponding to a 2 × 2 factorial design with Script (Kanji, Hiragana) and Visual Familiarity (High, Low) as factors.

The stimuli were carefully matched along several different dimensions summarized in Table 1a. These values, from the NTT database (Amano & Kondo, 2003a, 2003b), were analysed with a 2 × 2 ANOVA. Across conditions words were matched for mora length (a measure of phonological complexity). In addition, the analysis confirmed the main effect of Visual Familiarity and demonstrated that this did not interact with Script. Conceptual familiarity (derived by summing familiarity ratings for the visual word and for its auditory form) was matched across Script but naturally it was not possible to match across Visual Familiarity. Finally, Hiragana words had significantly more characters but fewer strokes than Kanji words, which is an inevitable difference between the scripts. It is worth noting, however, that effects of visual complexity and word length are expected to manifest in early visual cortices (Hsu, Lee, & Marantz, 2011; Mechelli, Humphreys, Mayall, Olson, & Price, 2000; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999) rather than in higher order visual regions like vOT.

The same stimuli were then re-grouped according to their lexical frequency – a script-independent measure of how often the word occurs in print regardless of its visual form (i.e. Kanji or Hiragana). Lexical frequency values were calculated by summing the frequency of the Kanji and Hiragana word forms, taken from the NTT database (Amano & Kondo, 2003a, 2003b). For example, the Japanese word pronounced /tembo:/ has a lexical frequency value of 6984, since its written Kanji form ("展望") has a frequency of 6979 and its Hiragana form ("てんぼう") has a frequency value of 5. The frequencies for the 60 lexical items were then divided into quartiles so that those within the lower quartile (i.e. low frequency words) could be compared to those within the upper quartiles (i.e. high frequency words) in order to maximize the distinction between them. Because the stimuli were originally chosen according to their visual familiarity scores across scripts, the distribution of the (log of the) lexical frequencies was nearly uniform over the 60 words. Note that it proved impossible to fully balance visual familiarity, script and lexical frequency into a factorial design, forcing us to interrogate the data in two separate analyses. In order to separate visual familiarity into high and low, each lexical item needed to be presented in both scripts. In contrast, lexical frequency was independent of script. As a result, a full factorial design that included Visual Familiarity, Lexical Frequency and Script was impossible to generate. There were no significant differences between high and low frequency items in terms of number of mora, number of characters, or total stroke count (Table 1b).

Finally, Kanji nonwords were created by combining random Kanji characters that together did not form a word. These were matched 1:1 with the real Kanji words for number of strokes and characters. Hiragana nonwords were created by combining random Hiragana characters that together did not form a word. These were matched 1:1 with the real Hiragana words for number of strokes and characters.

2.4. MRI acquisition

For the subjects scanned in Tokyo, whole-brain imaging was performed on a Siemens 3T MRI scanner at the Brain Science Research Center at Tamagawa University. The functional data were acquired with a gradient-echo EPI sequence (TR = 3000 ms; TE = 25 ms; FOV = 192 mm; matrix = 64 × 64) giving a notional resolution of 3 × 3 × 3. For participants in London, whole brain imaging was performed on a Siemens 1.5T MRI scanner at the Birkbeck-UCL Neuroimaging (BUCNI) centre. The functional data were

Table 1b
Stimulus matching for the lexical frequency analysis.

	Upper	Lower	t
Mora length	2.8	3.1	1.8 (n.s.)
Lexical frequency	12,100	25	4.5 ($p < 0.001$)
Visual familiarity	4.9	4.1	2.6 ($p = 0.011$)
Number of characters	2.5	2.6	0.7 (n.s.)
Number of strokes	14.0	13.1	0.4 (n.s.)

Table 1a
Stimulus matching for the visual familiarity analyses.

	Kanji		Hiragana		Effect of		
	High	Low	High	Low	Script	VisFam	Interaction
Mora length[a]	2.77	3.03	3.03	2.77	$t = 1.46$ ($p = 0.145$)		
Visual familiarity	5.31	3.66	5.21	4.09	0.9 (n.s.)	69.5 ($p < 0.001$)	2.3 (n.s.)
Combined familiarity	5.36	3.91	5.25	4.06	0.0 (n.s.)	64.5 ($p < 0.001$)	0.5 (n.s.)
Number of characters	2.00	2.07	3.07	2.90	77.7 ($p < 0.001$)	0.2 (n.s.)	1.2 (n.s.)
Number of strokes	19.17	19.70	7.60	6.63	217.0 ($p < 0.001$)	0.8 (n.s.)	0.8 (n.s.)

[a] Because the same words were used in the Kanji high and Hiragana low conditions as well as in the Kanji low and Hiragana high conditions, script-independent measures were the same for these conditions. Consequently, differences were assessed with a t-test rather than an ANOVA and are marked in italics.

acquired with a gradient-echo EPI sequence (TR = 3000 ms; TE = 50 ms; FOV = 192 mm; matrix = 64 × 64) giving a notional resolution of 3 × 3 × 3. In both cases, a run consisted of 186 volumes and as a result the two runs together took 18 min 24 s. In addition, a high-resolution (1 mm^3) T1-weighted anatomical scan was acquired for localizing the functional data on the individual's brain anatomy.

2.5. Analyses

In the both the behavioral and imaging data, items whose accuracy was at chance (⩽50%) were excluded from all analyses (n = 9) and only correct trials were analysed. Reaction times (RTs) were recorded from the onset of the stimulus and anticipatory responses (i.e. RTs < 300 ms) were trimmed (0.05% of trials). To minimize the effect of outliers, median RTs per condition per subject were used in the statistical analyses (Ulrich & Miller, 1994). The behavioral data were analysed using a mixed 2 × 2 × 2 analysis of variance (ANOVA) with Script (Kanji, Hiragana), Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor. Accuracy and RTs were the dependent measures. In addition, the behavioral data were then re-grouped into quartiles according to lexical frequency of the stimuli and analysed using a repeated-measures 4 × 2 × 2 ANOVA with Lexical Frequency (Upper, Upper Middle, Lower Middle, Lower) and Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor.

The imaging data were processed using SPM8 (Wellcome Trust Centre for Neuroimaging, London UK, http://www.fil.ion. ucl.ac.uk/spm/). The first four volumes in the Tokyo (i.e. 3T) data and two volumes in London (i.e. 1.5T) data were discarded in order to allow for T1 equilibrium. All functional volumes were spatially realigned and unwarped to adjust for minor distortions in the B0 field due to head movement (Andersson, Hutton, Ashburner, Turner, & Friston, 2001). They were then normalized to the MNI-152 EPI template, maintaining the original 3 × 3 × 3 mm resolution. Finally, images were smoothed with an isotropic 8 mm full-width half-maximum Gaussian kernel. Time-series from each voxel were high-pass filtered (1/128 Hz cutoff) to remove low-frequency noise and signal drift. The preprocessed functional volumes were then analysed in two separate GLMs. One investigated the effects of visual familiarity and script while the other investigated the effect of lexical frequency and visual familiarity. In both cases, a first-level, fixed-effects analysis combined the two runs from each participant and the estimated effect sizes were entered into a second-level, random-effects analysis to estimate the population effect. At the first level, the onsets of stimuli were modelled as delta functions convolved with a canonical haemodynamic response function (Glover, 1999), which provided regressors for the general linear model. The appropriate contrast images, averaged over sessions, were then generated in all subjects for each condition versus fixation.

The first analysis included four word conditions (Kanji high visual familiarity, Kanji low visual familiarity, Hiragana high visual familiarity and Hiragana low visual familiarity), two nonwords conditions (Kanji, Hiragana) and a condition for incorrect and excluded trials (Murphy & Garavan, 2004). Fixation was not modelled and served as an implicit baseline. The four word-relative-to-rest contrasts were computed and entered into a second-level, 2 × 2 × 2 ANOVA with Script (Kanji, Hiragana), Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor. We first identified areas of common activations for all eight word conditions using a linear contrast to compute their mean activity and inclusively masking it with each condition relative to fixation at p = 0.001. From this analysis we

computed statistical contrasts of the two conditions within a factor, inclusively masking them with common activations of these conditions at p = 0.05.

The second analysis investigated the effect of lexical frequency independent of script. Here, words were divided into quartiles based on their frequency and entered into a second-level, 4 × 2 × 2 ANOVA with Lexical Frequency (Upper, Upper middle, Lower middle, Lower) and Visual Familiarity (High, Low) as within-subject factors and Group (Tokyo, London) as a between-subject factor.

Since the primary aim of the current study was to investigate effects in vOT, we defined an *a priori* anatomical mask for this region-of-interest (ROI). The main anatomical areas of interest were the occipitotemporal sulcus and adjacent regions on the crests of the fusiform and inferior temporal gyri: areas consistently activated by visual word recognition tasks (Bitan et al., 2007; Cai, Paulignan, Brysbaert, Ibarrola, & Nazir, 2010; Cohen et al., 2000; Devlin, Jamison, Gonnerman, & Matthews, 2006; Duncan, Pattamadilok, Knierim, & Devlin, 2009; Fiez & Petersen, 1998; Frost et al., 2005; Herbster, Mintun, Nebes, & Becker, 1997; Kronbichler et al., 2007; Price, Wise, & Frackowiak, 1996; Rumsey et al., 1997; Shaywitz et al., 2004). Because the precise coordinates vary along a rostro-caudal axis, standard space coordinates ranging from $X = -36$ to -54 and $Y = -45$ to -66 were used to delineate this region. In addition, the depth of the sulcus coupled with the fact the temporal lobe is angled downwards required a range of Z-coordinates as well ($Z = -30$ to -6). Together these coordinates describe a rectangular prism that conservatively encompassed the region of vOT sensitive to visual word recognition as well as the anatomically adjacent lobule VI of the cerebellum. Because the cerebellum was both anatomically and functionally distinct, it was manually removed from the ROI mask.

For all imaging analyses, activations were considered significant based on voxel-level inference of $p < 0.05$, corrected for multiple comparisons either within the ROI ($Z > 3.30$) or across the entire brain ($Z > 4.60$). In order to visualize the pattern of activation within a region, we plotted the mean effect size per condition within a 5 mm-radius sphere centered on the peak coordinate. No inferential statistics were based on these effect size plots.

3. Results

The behavioral results are illustrated in Fig. 1, where the left panel displays accuracy scores and the right reaction times. Within the accuracy data, there was a significant main effect of Visual Familiarity ($F(1,33) = 32.8$, $p < 0.001$), confirming that less visually familiar word forms were more difficult. This was qualified by a significant Visual Familiarity × Script interaction ($F(1,33) = 26.6$, $p < 0.001$), indicating that the visual familiarity advantage was significant for Kanji ($t(33) = 6.6$, $p < 0.001$) but not for Hiragana ($t(33) = 0.9$, $p = 0.365$, n.s.). In addition, there was a significant main effect of Script ($F(1,33) = 14.5$, $p = 0.001$) indicating responses to Hiragana were more accurate than to Kanji. The analysis of the reaction time data revealed a similar pattern of results. There was a main effect of Visual Familiarity ($F(1,33) = 37.5$, $p < 0.001$) with responses to less visually familiar forms taking longer than those to highly familiar forms (854 vs. 775 ms). This was qualified by a significant interaction ($F(1,33) = 10.1$, $p = 0.003$) indicating that familiarity effect was larger for Kanji (118 ms) than Hiragana (39 ms). The main effect of Script was not significant ($F(1,33) = 2.7$, $p = 0.106$).

The second behavioral analysis focused on lexical frequency and visual familiarity and the data are shown in Fig. 2. There was a main effect of Lexical Frequency for both accuracy ($F(3,99) = 35.9$, $p < 0.001$) and RTs ($F(3,99) = 24.4$, $p < 0.001$). From the
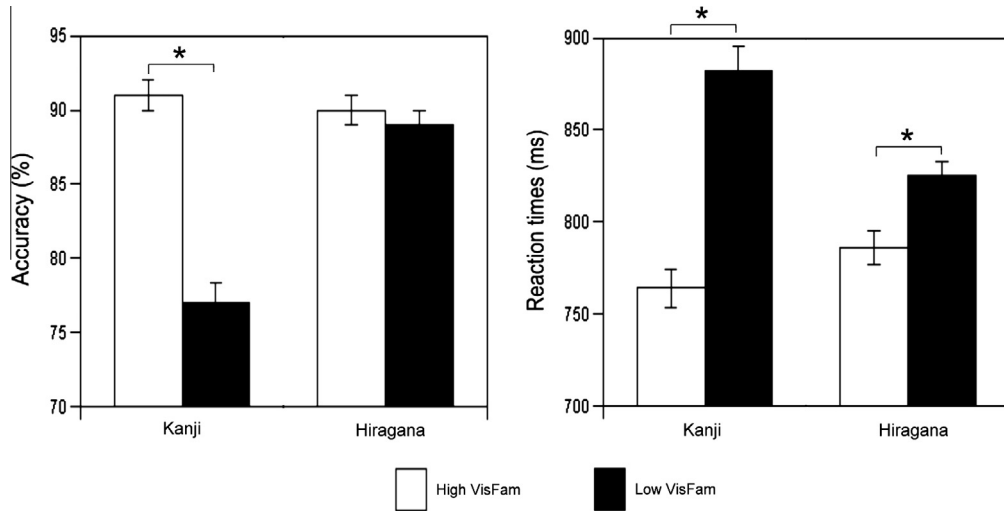
Fig. 1. Mean accuracy (left panel) and reaction times (right panel) for four conditions. An asterisk indicates significant at $p = 0.05$.
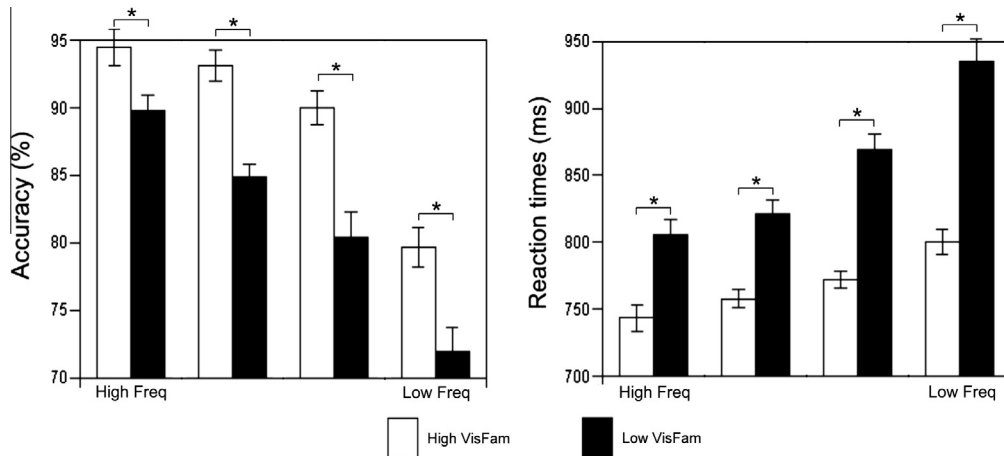

Fig. 2. Mean accuracy (left panel) and reaction times (right panel) for the Lexical Frequency × Visual Familiarity analysis. An asterisk indicates significant at $p = 0.05$.

figure it is clear that the lower the lexical frequency, the more difficult the word was with both lower accuracy and longer response times. There was also a main effect of Visual Familiarity for both accuracy ($F(1,33) = 42.5$, $p < 0.001$) and RTs ($F(1,33) = 88.9$, $p < 0.001$). In addition, there was a significant interaction for RTs ($F(3,99) = 5.5$, $p = 0.002$), indicating that visual familiarity had a greater effect on low relative to high frequency items although the interaction for accuracy was not significant ($F(3,99) = 1.4$, $p = 0.257$). This pattern of results remained exactly the same when only the Upper and Lower frequency quartiles were included in the ANOVA. In other words, both lexical frequency and visual familiarity strongly affected overall performance of the task. The question then becomes: to what extent do these two factors affect activation in vOT?

3.1. Imaging

The first imaging analysis identified brain regions commonly activated by all four word conditions, in order to determine whether vOT (among other regions) was engaged by both Kanji and Hiragana words, independent of their visual familiarity. As expected, vOT was strongly activated bilaterally, centered on the pos-

terior occipitotemporal sulcus and extending laterally into inferior temporal gyrus, medially onto the crest of the fusiform gyrus and inferiorly into lobule VI of the cerebellum. Other bilateral activations included pars opercularis, the pre-supplemental motor area (pre-SMA), the intraparietal sulcus, a mid-cingulate region, and parts of the basal ganglia. In addition, there were several activations only seen in the left hemisphere including Broca's complex (i.e. pars triangularis, pars orbitalis), the deep frontal operculum, the supramarginal gyrus, and a small cluster in the anterior fusiform gyrus (see Table 2 for full details). These results are consistent with previous visual word recognition studies conducted in alphabetic (Carreiras, Mechelli, Estevez, & Price, 2007; Devlin et al., 2006; Fiebach, Ricker, Friederici, & Jacobs, 2007; Hauk et al., 2008) and logographic (Booth et al., 2006; Chen, Fu, Iversen, Smith, & Matthews, 2002; Fu, Chen, Smith, Iversen, & Matthews, 2002; Hu et al., 2010; Kuo et al., 2003; Tan et al., 2001) languages, indicating a common system engaged by visual word processing, independent of script.

Next, we asked whether visual familiarity modulated vOT activation. The comparison of low relative to high visual familiarity items revealed significant left vOT activation in the ROI [−45 −58 −11, $Z = 3.34$, $p = 0.042$] and this did not interact with Script

Table 2
Common activation across the four conditions relative to fixation. For each region, the standard space (MNI) coordinates of the peak voxel and the Z-score for the main effect of words relative to rest at that voxel are shown. In addition, the final four columns display the Z-score for each condition relative to rest at that same coordinate.

Region		Mean peak coordinate				Z relative to rest			
		x	y	z	Z-score	Kanji Hi	Kanji Lo	Hira Hi	Hira Lo
	Frontal								
L	Pars opercularis	−42	8	31	8.19	6.55	7.99	7.80	7.96
R	Pars opercularis	45	5	31	6.84	4.22	6.49	5.94	5.18
L	Pars orbitalis	−39	35	1	6.92	3.19	5.87	6.23	6.88
L	Pars triangularis	−48	29	19	7.91	5.85	7.41	7.52	7.86
L	Frontal operculum	−33	20	1	7.76	4.29	6.74	6.38	7.25
L	Pre-SMA	−6	14	52	7.86	6.41	7.35	7.35	7.69
R	Pre-SMA	6	14	52	6.39	3.59	5.27	5.11	6.57
R	Cingulate gyrus	6	−7	28	5.41	3.73	4.41	3.96	5.14
L	Cingulate gyrus	−6	−7	28	5.24	3.39	4.58	3.96	4.76
	Parietal								
R	IPS	33	−58	43	6.73	4.61	5.70	5.57	5.83
L	IPS	−30	−58	43	7.84	6.31	7.43	7.20	7.63
L	SMG	−47	−37	49	7.63	5.81	6.38	7.19	7.67
	Occipital/temporal								
L	vOT	−45	−64	−14	8.46	8.12	7.99	7.88	8.17
R	vOT	44	−55	−20	7.42	7.09	6.51	5.65	6.76
L	ITG	−60	−43	−14	5.56	4.04	4.47	5.84	5.00
R	ITG	63	−40	−14	4.82	3.41	3.41	4.03	4.61
L	Ant. fusiform gyrus	−39	−28	−26	5.62	5.17	5.21	3.44	3.78
	Subcortical								
R	Cerebellum (VI)	33	−55	−26	8.17	7.81	7.33	7.58	7.98
L	Cerebellum (VI)	−39	−52	−26	8.01	7.67	7.46	7.27	7.48
R	Globus pallidus	12	2	−2	5.40	4.16	4.32	4.37	4.41
L	Putaman	−21	5	4	7.69	5.49	6.02	6.41	7.01
R	Putaman	21	5	−2	5.06	3.37	3.55	4.31	5.05
L	Thalamus	−9	−13	7	6.76	3.89	5.79	5.77	6.35
R	Caudate nucleus (body)	21	−10	22	5.16	3.36	5.10	3.19	4.63
L	Caudate nucleus (head)	−15	8	13	5.50	3.53	4.46	4.02	5.51
R	Caudate nucleus (head)	15	8	13	4.91	3.77	3.62	3.29	4.99

Abbreviations: IPS = intra-parietal sulcus; ITG = inferior temporal gyrus; SMA = supplemental motor area; SMG = supramarginal gyrus; vOT = ventral occipito-temporal cortex.

(Fig. 3a and b). Outside of the ROI, the whole brain search revealed activation in left pars triangularis, left pars opercularis, and in the frontal operculum bilaterally (Table 3) – three regions previously associated with low relative to high frequency effects in alphabetic languages (Carreiras, Mechelli, & Price, 2006; Fiebach et al., 2002; Kronbichler et al., 2004, 2007). In each of these regions, there was a (non-significant) interaction with Script such that low visual familiarity items increased activation for Kanji more than for Hiragana relative to high visual familiarity items. No regions showed significant activation for the contrast of high relative to low visual familiarity, even when the statistical threshold was lowered to $p < 0.001$, uncorrected for multiple comparisons.

Next we turned to the effects of lexical frequency. We contrasted the lowest frequency items to the highest in order to maximize the difference in frequency. Activation associated with lexical frequency was found in several left frontal regions, including pars opercularis, pars triangularis, a region of anterior paracingulate sulcus and the right deep frontal operculum (Table 4), consistent with previous studies (Carreiras et al., 2006; Fiebach et al., 2002; Hauk et al., 2008; Kronbichler et al., 2004). In addition, the ROI analysis identified a peak in lateral inferior temporal gyrus [−54 −55 −14, $Z = 4.15$, $p = 0.003$] adjacent to, but not overlapping, the activation seen in vOT for visual familiarity (Fig. 3c). In fact it was approximately 1 cm lateral to the visual familiarity peak and located in the inferior temporal gyrus, rather than the occipitotemporal sulcus.

Within vOT, there was no main effect of frequency (Fig. 3d). There was, however, a small peak for the interaction of frequency and visual familiarity at [−45 −64 −8], although this did not reach statistical reliability ($Z = 2.81$, $p = 0.177$). Nonetheless, the pattern

of activation across conditions suggests that visual familiarity modulated low frequency words but not high frequency words (see Fig. 3e). Words in the middle frequency quartiles showed intermediate sized visual familiarity effects.

Finally, we turned to the question of whether the different scripts, Kanji and Hiragana, influenced vOT activation. Relative to Hiragana, Kanji produced significantly greater activation within vOT (Fig. 3a and b). The peak was slightly posterior to the visual familiarity peak, although the clusters of activation were largely overlapping (Fig. 3a). Outside the vOT region-of-interest no significant activation was found in the whole brain search. The opposite comparison of Hiragana relative to Kanji revealed no significant activation.

To investigate whether the increased activation in vOT for Kanji relative to Hiragana effect was driven by the inevitable difference in the visual complexity across scripts, a $2 \times 2 \times 2$ ANOVA with Script (Kanji, Hirgana), Visual Complexity (high, low), and Group (Tokyo, London) was run. Since the two scripts differed in both the number of strokes and of characters, the total number of strokes per trial was used as the measure of visual complexity. There was no main effect of Visual Complexity nor interaction between Visual Complexity and the other two factors within vOT, even at a lenient statistical threshold of $p < 0.001$ uncorrected for multiple comparisons. In other words, there was no evidence that the vOT activation observed for Kanji relative to Hiragana was a by-product of the greater visual complexity of Kanji words. Outside of our region-of-interest, high relative to low visual complexity was associated with activation in the left calcarine sulcus [−15 −94 −5, $Z = 3.68$] and the right lingual gyrus [18 −91 −20, $Z = 3.90$] at a threshold of $p = 0.001$ (uncorrected).
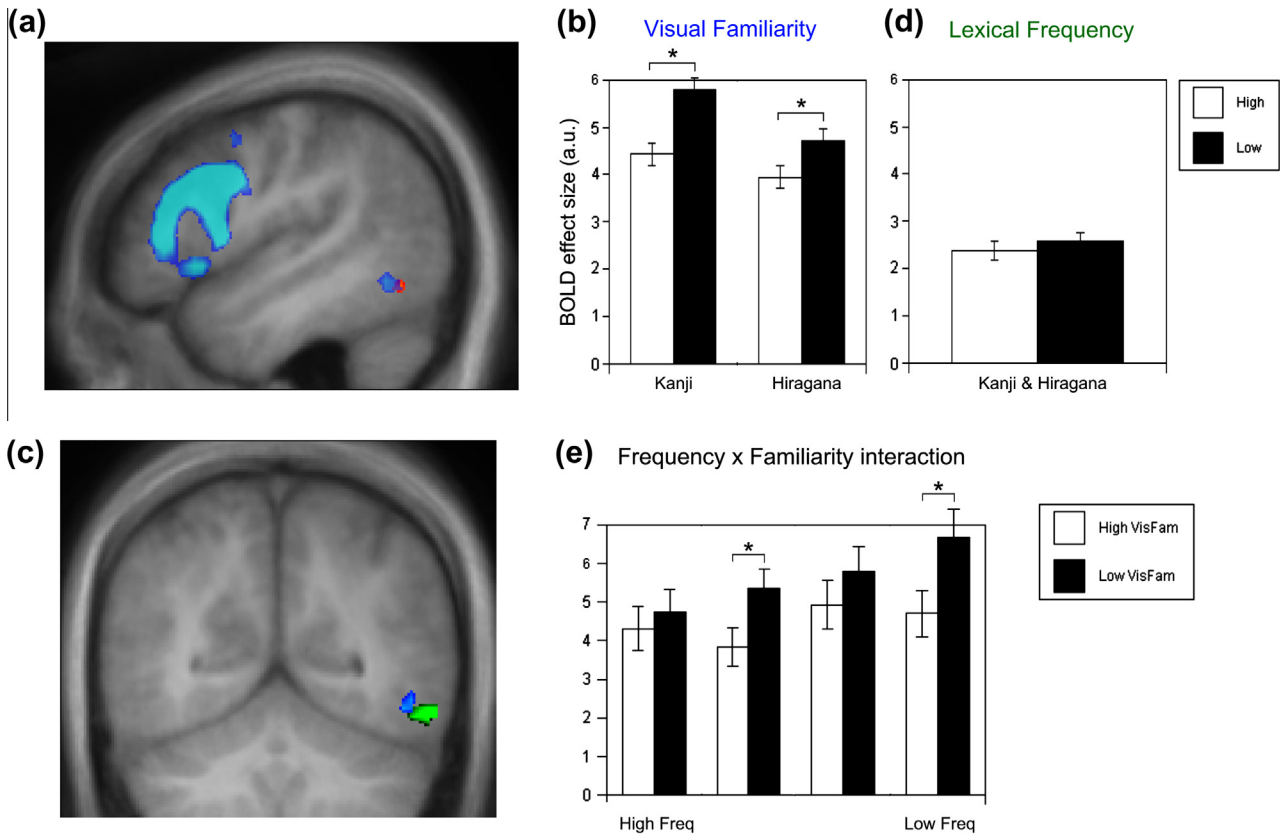
Fig. 3. (a) The overlap (purple) of Visual Familiarity (low > high) in blue and Script (Kanji > Hiragana) in red at $x = -48$. (b) Bar plot for the BOLD signal per condition relative to fixation at $x = -45$, $y = -58$, $z = -11$, showing Visual Familiarity effect. (c) Visual Familiarity effect (low > high) in blue and Lexical Frequency effect (Lower > Upper) in green at $y = -58$. Also shown are bar plots for the BOLD signal per condition relative to fixation at $x = -45$, $y = -58$, $z = -11$ for (d) Lexical Frequency and (e) Frequency × Familiarity interaction.

Table 3
Activation for low relative to high visual familiarity.

Region		Mean peak coordinate			Z-score			
		x	y	z	Low > high	Kanji low > high	Hiragana low > high	VisFam × script
L	vOT	−45	−58	−11	3.34	2.80	1.96	0.95
L	Pars opercularis	−42	5	31	5.22	5.54	1.61	3.48
L	Pars triangularis	−42	29	7	4.92	5.00	1.82	2.87
L	Frontal operculum	−30	29	1	5.06	4.72	2.45	2.24
R	Frontal operculum	30	20	−5	5.38	5.56	1.89	3.34

Table 4
Activation for lower > upper frequency.

Region		Mean peak coordinate			
		x	y	z	Z-score
L	Pars triangularis	−45	26	19	4.65
L	Pars opercularis	−45	11	28	4.74
L	Inferior frontal sulcus	−51	17	34	4.74
L	Paracingulate sulcus	−3	26	46	6.05
R	Frontal operculum	30	23	−5	4.67

4. Discussion

The aim of the current study was to test whether visual familiarity and lexical frequency have separable effects on activation levels in vOT, as predicted by the Interactive Account. The results confirmed that visual familiarity, as opposed to lexical frequency, had a strong effect on vOT activation that was qualified by a small (but non-significant) interaction. Visual familiarity had essentially no effect on the most frequent words but a greater effect on the least frequent. In contrast, lexical frequency modulated activation in a region of the inferior temporal gyrus lateral to the visual familiarity effect in vOT. Finally, vOT also showed higher activation for Kanji than Hiragana words, although this was not due to their inherent differences in visual complexity. These findings place important constraints on understanding the nature of neural information processing in the region.

Given that vOT is a region of extrastriate visual cortex, it is perhaps not surprising that the region is sensitive to the familiarity of visual patterns. Indeed, a visual familiarity effect for faces in vOT has been previously reported (Eger, Schweinberger, Dolan, & Henson, 2005). Although written words are a special form of familiar visual patterns, they too appear to be sensitive to this basic property of the visual system (Nazir, Ben-Boutayab, Decoppet, Deutsch, & Frost, 2004; Xue, Chen, Jin, & Dong, 2006; Xue, Jiang, Chen, & Dong, 2008). Within a predictive coding account, this is implemented in terms of more accurate top-down predictions for highly familiar visual patterns. This reduces the prediction error between

the top-down and the bottom-up signals, reducing the regional BOLD signal. It is worth noting, however, that this visual familiarity effect is likely to be task-dependent and only present in tasks that place strong demands on integrating bottom-up visual form information with top-down non-visual properties of the stimulus. In the current experiment, linking visual forms to their sound and meaning is important for either recognizing them as a word or correctly rejecting them as a nonword. Tasks with similar demands on vOT processing such as reading aloud or reading for meaning would also be expected to demonstrate greater vOT activation for less visually familiar words. In contrast, one-back tasks or purely perceptual decisions may not show a visual familiarity effect for written words because neither places significant demands on integrating bottom-up visual and top-down non-visual information (Hellyer, Woodhead, Leech, & Wise, 2011; Price & Devlin, 2011; Wang, Yang, Shu, & Zevin, 2011).

Like visual familiarity, lexical frequency would also be expected to modulate the accuracy of top-down predictions into vOT. Indeed, in alphabetic languages, where there is essentially a single visual pattern per word, lexical frequency does affect vOT activation (Chee, Westphal, et al., 2003; Hauk et al., 2008; Joubert et al., 2004; Kronbichler et al., 2004). In Japanese, however, where a word can be written in different scripts, there was no significant main effect of lexical frequency on vOT activation and only weak evidence of its interaction with visual familiarity. For highly frequent words, visual familiarity had no effect on vOT activation whereas activation for low frequency words was modulated by visual familiarity. In contrast, lexical frequency was found to significantly modulate activation in a region of the left inferior temporal gyrus lateral to the area in vOT showing a visual familiarity effect. Previous studies have argued that this is a functionally separate region engaged by multi-modal semantic processing rather than by visual word forms (Cohen, Jobert, Le Bihan, & Dehaene, 2004; Moore & Price, 1999) and our results are consistent with this.

Finally, although both Kanji and Hiragana strongly engaged vOT, there was significantly greater activation for Kanji relative to Hiragana. Because Kanji words generally have more strokes and fewer characters, they tend to be more visually complex than words written in Hiragana. Supplemental analyses, however, showed that the effects of visual complexity manifested in early visual cortices rather than in vOT, consistent with most previous studies (Hsu et al., 2011; Tarkiainen et al., 1999 but see Szwed et al., 2011 who failed to find any significant effects of visual complexity). Thus, we assume the activation difference in vOT reflects the different links between the surface form of the word and its non-visual properties. Specifically, the relation between a Kanji word and its phonological form is largely arbitrary and depends critically on the combination of characters present in the word. Hiragana characters, on the other hand, are nearly 100% consistent in their pronunciation with each one representing a one-to-one mapping to a mora. Consequently, the contribution to prediction error from phonological regions to vOT is much less for Hiragana than for Kanji, resulting in lower vOT activation, analogous to the effect of the orthographic transparency in reading (Nosarti, Mechelli, Green, & Price, 2010; Paulesu et al., 2000).

According to the Interactive Account (Price & Devlin, 2011), all words generate top-down predictions that arrive at vOT and support visual forms consistent with the word (Devlin et al., 2006; Kherif, Josse, & Price, 2011). For instance, a word such as "yen" sends top-down predictions to vOT that support its Kanji (円), Hiragana (えん), and symbolic (¥) form. These predictions are highly accurate precisely because the word is so common, resulting in essentially equal activation across scripts. In contrast, a less common word such as "wit" sends less accurate top-down predictions to vOT supporting both its Hiragana (とんち) and Kanji (頓知) forms. In this case, the fact that Hiragana is the more visually familiar

form results in less prediction error (and therefore less activation) than the Kanji form. By this account, then, a visual familiarity benefit is principally expected for low, but not highly, frequent words – a pattern demonstrated in Fig. 3e, but only weakly. Further studies will be required to establish whether this prediction of an interaction between visual familiarity and lexical frequency in vOT is reliable.

Can the current results also be understood in terms of orthographic input lexicon accounts that posit specialized representations of whole word orthographic patterns? In its strongest form, entries in the lexicon are truly "lexical" and abstract away from visual properties of the word such as capitalization, font, size and even script. Consequently there is a single lexical entry for a word regardless of its script. Obviously, this version is incompatible with the current findings because it cannot explain the activation differences in vOT due to visual familiarity nor script. Some authors have argued, however, that lexical entries in an orthographic input lexicon are specific not only to words but also letter or case identities (Kronbichler et al., 2009). In the case of Japanese orthography, this would entail separate entries for the Kanji and Hiragana forms. By this account, the ease of accessing the form would be modulated by one's experience with the pattern (i.e. its visual familiarity), and is consistent with the main effect of visual familiarity observed here. The observed differences in activation for Kanji and Hiragana, however, are problematic for accounts that claim vOT is an orthographic input lexicon (Glezer et al., 2009; Kronbichler et al., 2004). If every word in an orthographic lexicon has separate Kanji and Hiragana entries and the accessibility of each entry depends on its visual familiarity, then Hiragana and Kanji should produce equivalent activation unless they differ in terms of visual familiarity. Indeed, some previous experiments that reported greater vOT activation for Kanji relative to Hiragana may have confounded script with visual familiarity (Ha Duy Thuy et al., 2004; Nakamura et al., 2005; Sakurai et al., 2000), but in the current experiment, visual familiarity was carefully balanced across scripts so that Kanji was no more familiar than Hiragana (see also Ino et al., 2009). Nonetheless, we still observed a significant increase in activation for Kanji relative to Hiragana that may prove a challenge for the orthographic input lexicon account.

The current findings also raise an important methodological point about the use of alphabetic scripts in reading research. For many valid reasons, the majority of reading research has been conducted with alphabetic languages and has produced considerable advances in the cognitive and neural mechanisms underlying reading (Price & Mechelli, 2005; Shaywitz et al., 2004; Ziegler & Goswami, 2005). In addition, it has repeatedly been shown that both alphabetic and non-alphabetic scripts such as logographs engage essentially the same neuroanatomical system during reading (Booth et al., 2006; Chee, Soon, & Lee, 2003; Chee et al., 2000; Chen et al., 2002; Fu et al., 2002; Hu et al., 2010; Kuo et al., 2003; Tan et al., 2001). This finding, however, means that inferences drawn regarding the nature of regional information processing need to be consistent with a range of writing systems (i.e. alphabetic, syllabographic, logographic, etc.) in order to explain the common information processing across scripts. For instance, Chinese readers engage a region of left middle frontal gyrus (MFG) more strongly than English readers (Tan, Laird, Li, & Fox, 2005; Tan et al., 2003) and it is theoretically possible that this is due to language-specific neuronal responses. A more parsimonious explanation, however, is that Hanji characters increase visuospatial working memory demands relative to letters and this explanation has the advantage of being consistent with reports of left MFG activation in non-linguistic studies of visual working memory (Bledowski, Kaiser, & Rahm, 2010; Kravitz, Saleem, Baker, & Mishkin, 2011). Moreover, it offers a principled account for why this differential activation disappears in Chinese and English dyslexics (Hu et al., 2010). In

short, a unified, cross-cultural account of the neural information processing underlying reading requires a systematic investigation of a range of different languages and scripts.

The current study took advantage of a unique property of the Japanese writing system in order to better characterize neural information processing in vOT during visual word recognition. Here we demonstrate a dissociation between lexical frequency and visual familiarity not possible in alphabetic languages and use the findings to evaluate competing theories of vOT function.

Acknowledgments

References

Amano, S., & Kondo, T. (2003a). Nihongo-no Goi-tokusei [Lexical properties of Japanese] (Vol. 7), CD-ROM version (in Japanese), NTT database series. Tokyo Sanseido.

Amano, S., & Kondo, T. (2003b). Nihongo-no Goi-tokusei [Lexical properties of Japanese] (Vols. 1–6), CD-ROM version (in Japanese), NTT database series. Tokyo: Sanseido.

Andersson, J. L., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. NeuroImage, 13(5), 903–919.

Bitan, T., Cheon, J., Lu, D., Burman, D. D., Gitelman, D. R., Mesulam, M. M., et al. (2007). Developmental changes in activation and effective connectivity in phonological processing. NeuroImage, 38(3), 564–575.

Bledowski, C., Kaiser, J., & Rahm, B. (2010). Basic operations in working memory: Contributions from functional imaging studies. Behavioural Brain Research, 214(2), 172–179.

Booth, J. R., Lu, D., Burman, D. D., Chou, T. L., Jin, Z., Peng, D. L., et al. (2006). Specialization of phonological and semantic processing in Chinese word reading. Brain Research, 1071(1), 197–207.

Bruno, J. L., Zumberge, A., Manis, F. R., Lu, Z. L., & Goldman, J. G. (2008). Sensitivity to orthographic familiarity in the occipito-temporal region. NeuroImage, 39(4), 1988–2001.

Cai, Q., Paulignan, Y., Brysbaert, M., Ibarrola, D., & Nazir, T. A. (2010). The left ventral occipito-temporal response to words depends on language lateralization but not on visual familiarity. Cerebral Cortex, 20(5), 1153–1163.

Carreiras, M., Mechelli, A., Estevez, A., & Price, C. J. (2007). Brain activation for lexical decision and reading aloud: Two sides of the same coin? Journal of Cognitive Neuroscience, 19(3), 433–444.

Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. Human Brain Mapping, 27(12), 963–972.

Chee, M. W. L., Soon, C. S., & Lee, H. L. (2003). Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation. Journal of Cognitive Neuroscience, 15(1), 85–97.

Chee, M. W. L., Weekes, B., Lee, K. M., Soon, C. S., Schreiber, A., Hoon, J. J., et al. (2000). Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: Evidence from fMRI. NeuroImage, 12(4), 392–403.

Chee, M. W. L., Westphal, C., Goh, J., Graham, S., & Song, A. W. (2003). Word frequency and subsequent memory effects studied using event-related fMRI. NeuroImage, 20(2), 1042–1051.

Chen, Y., Fu, S., Iversen, S. D., Smith, S. M., & Matthews, P. M. (2002). Testing for dual brain processing routes in reading: A direct contrast of Chinese character and Pinyin reading using fMRI. Journal of Cognitive Neuroscience, 14(7), 1088–1098.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (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(Pt 2), 291–307.

Cohen, L., Jobert, A., Le Bihan, D., & Dehaene, S. (2004). Distinct unimodal and multimodal regions for word processing in the left temporal cortex. NeuroImage, 23(4), 1256–1270.

Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006). The role of the posterior fusiform gyrus in reading. Journal of Cognitive Neuroscience, 18(6), 911–922.

Duncan, K. J., Pattamadilok, C., Knierim, I., & Devlin, J. T. (2009). Consistency and variability in functional localisers. NeuroImage, 46(4), 1018–1026.

Eger, E., Schweinberger, S. R., Dolan, R. J., & Henson, R. N. (2005). Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. NeuroImage, 26(4), 1128–1139.

Fiebach, C. J., Friederici, A. D., Müller, K., & Von Cramon, D. Y. (2002). FMRI evidence for dual routes to the mental lexicon in visual word recognition. Journal of Cognitive Neuroscience, 14(1), 11–23.

Fiebach, C. J., Ricker, B., Friederici, A. D., & Jacobs, A. M. (2007). Inhibition and facilitation in visual word recognition: Prefrontal contribution to the orthographic neighborhood size effect. NeuroImage, 36(3), 901–911.

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. Proceedings of the National Academy of Sciences of the United States of America, 95(3), 914–921.

Frost, S. J., Mencl, W. E., Sandak, R., Moore, D. L., Rueckl, J. G., Katz, L., et al. (2005). A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud. NeuroReport, 16(6), 621–624.

Fu, S., Chen, Y., Smith, S., Iversen, S., & Matthews, P. M. (2002). Effects of word form on brain processing of written chinese. NeuroImage, 17(3), 1538–1548.

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "Visual Word Form Area". Neuron, 62(2), 199–204.

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage, 9(4), 416–429.

Ha Duy Thuy, D., Matsuo, K., Nakamura, K., Toma, K., Oga, T., Nakai, T., et al. (2004). Implicit and explicit processing of kanji and kana words and non-words studied with fMRI. NeuroImage, 23(3), 878–889.

Hauk, O., Davis, M. H., & Pulvermüller, F. (2008). Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. NeuroImage, 42(3), 1185–1195.

Hellyer, P. J., Woodhead, Z. V. J., Leech, R., & Wise, R. J. S. (2011). An investigation of Twenty/20 vision in reading. Journal of Neuroscience, 31(41), 14631–14638.

Herbster, A. N., Mintun, M. A., Nebes, R. D., & Becker, J. T. (1997). Regional cerebral blood flow during word and nonword reading. Human Brain Mapping, 5(2), 84–92.

Hsu, C. H., Lee, C. Y., & Marantz, A. (2011). Effects of visual complexity and sublexical information in the occipitotemporal cortex in the reading of Chinese phonograms: A single-trial analysis with MEG. Brain and Language, 117(1), 1–11.

Hu, W., Lee, H. L., Zhang, Q., Liu, T., Geng, L. B., Seghier, M. L., et al. (2010). Developmental dyslexia in Chinese and English populations: Dissociating the effect of dyslexia from language differences. Brain, 133(6), 1694–1706.

Ino, T., Nakai, R., Azuma, T., Kimura, T., & Fukuyama, H. (2009). Recognition and reading aloud of kana and kanji word: An fMRI study. Brain Research Bulletin, 78(4–5), 232–239.

Joubert, S., Beauregard, M., Walter, N., Bourgouin, P., Beaudoin, G., Leroux, J. M., et al. (2004). Neural correlates of lexical and sublexical processes in reading. Brain and Language, 89(1), 9–20.

Kherif, F., Josse, G., & Price, C. J. (2011). Automatic top-down processing explains common left occipito-temporal responses to visual words and objects. Cerebral Cortex, 21(1), 103–114.

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. Nature Reviews Neuroscience, 12(4), 217–230.

Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., et al. (2007). Taxi vs. taksi: On orthographic word recognition in the left ventral occipitotemporal cortex. Journal of Cognitive Neuroscience, 19(10), 1584–1594.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. NeuroImage, 21(3), 946–953.

Kronbichler, M., Klackl, J., Richlan, F., Schurz, M., Staffen, W., Ladurner, G., et al. (2009). On the functional neuroanatomy of visual word processing: Effects of case and letter deviance. Journal of Cognitive Neuroscience, 21(2), 222–229.

Kuo, W. J., Yeh, T. C., Lee, C. Y., Wu, Y. T., Chou, C. C., Ho, L. T., et al. (2003). Frequency effects of Chinese character processing in the brain: An event-related fMRI study. NeuroImage, 18(3), 720–730.

Lee, C. Y., Tsai, J. L., Kuo, W. J., Yeh, T. C., Wu, Y. T., Ho, L. T., et al. (2004). Neuronal correlates of consistency and frequency effects on Chinese character naming: An event-related fMRI study. NeuroImage, 23(4), 1235–1245.

Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A., & Price, C. J. (2000). Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. Proceedings of the Royal Society B: Biological Sciences, 267(1455), 1909–1913.

Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipitotemporal regions for reading and object naming. NeuroImage, 10(2), 181–192.

Morton, J. (1969). Interaction of information in word recognition. Psychological Review, 76(2), 165–178.

Murphy, K., & Garavan, H. (2004). Artifactual fMRI group and condition differences driven by performance confounds. NeuroImage, 21(1), 219–228.

Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., & Kouider, S. (2005). Subliminal convergence of Kanji and Kana words: Further evidence for functional parcellation of the posterior temporal cortex in visual word perception. Journal of Cognitive Neuroscience, 17(6), 954–968.

Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. Brain and Language, 88(3), 294–311.

Nosarti, C., Mechelli, A., Green, D. W., & Price, C. J. (2010). The impact of second language learning on semantic and nonsemantic first language reading. Cerebral Cortex, 20(2), 315–327.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia, 9(1), 97–113.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al. (2000). A cultural effect on brain function. Nature Neuroscience, 3(1), 91–96.

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. Trends in Cognitive Sciences.

Price, C. J., & Mechelli, A. (2005). Reading and reading disturbance. Current Opinion in Neurobiology, 15(2), 231–238.

Price, C. J., Wise, R. J., & Frackowiak, R. S. (1996). Demonstrating the implicit processing of visually presented words and pseudowords. *Cerebral Cortex, 6*(1), 62–70.

Rogers, H. (2005). *Writing Systems: A Linguistic approach*. Malden, MA: Blackwell Publishing.

Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K., Maisog, J. M., & Andreason, P. (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain, 120*(5), 739–759.

Sakurai, Y., Momose, T., Iwata, M., Sudo, Y., Ohtomo, K., & Kanazawa, I. (2000). Different cortical activity in reading of Kanji words, Kana words and Kana nonwords. *Cognitive Brain Research, 9*(1), 111–115.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry, 55*(9), 926–933.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., et al. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage, 56*(1), 330–344.

Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping, 25*(1), 83–91.

Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *NeuroImage, 13*(5), 836–846.

Tan, L. H., Spinks, J. A., Feng, C. M., Siok, W. T., Perfetti, C. A., Xiong, J., et al. (2003). Neural systems of second language reading are shaped by native language. *Human Brain Mapping, 18*(3), 158–166.

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain, 122*(11), 2119–2131.

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General, 123*(1), 34–80.

van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., et al. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *NeuroImage, 47*(4), 1940–1949.

Wang, X., Yang, J., Shu, H., & Zevin, J. D. (2011). Left fusiform BOLD responses are inversely related to word-likeness in a one-back task. *NeuroImage, 55*(3), 1346–1356.

Xue, G., Chen, C., Jin, Z., & Dong, Q. (2006). Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. *NeuroImage, 31*(3), 1315–1326.

Xue, G., Jiang, T., Chen, C., & Dong, Q. (2008). Language experience shapes early electrophysiological responses to visual stimuli: The effects of writing system, stimulus length, and presentation duration. *NeuroImage, 39*(4), 2025–2037.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3–29.