Journal of Neurolinguistics 37 (2016) 58–67

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Rapid modulation of spoken word recognition by visual primes

Kana Okano [a, b, *], Jonathan Grainger [c, d], Phillip J. Holcomb [b, e]

[a] *Gabrieli Lab, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA*
[b] *Neurocognition Lab, Department of Psychology, Tufts University, Medford, MA, USA*
[c] *Laboratoire de Psychologie Cognitive, CNRS & Aix-Marseille University, Marseille, France*
[d] *Brain and Language Research Institute, Aix-Marseille University, France*
[e] *Department of Psychology, San Diego State University, San Diego, CA, USA*

A R T I C L E I N F O

Article history:
Received 14 November 2014
Received in revised form 4 September 2015
Accepted 8 September 2015
Available online 25 September 2015

Keywords:
Spoken word recognition
Cross-modal interactions
Japanese kana
ERPs
Masked priming

A B S T R A C T

In a masked cross-modal priming experiment with ERP recordings, spoken Japanese words were primed with words written in one of the two syllabary scripts of Japanese. An early priming effect, peaking at around 200 ms after onset of the spoken word target, was seen in left lateral electrode sites for Katakana primes, and later effects were seen for both Hiragana and Katakana primes on the N400 ERP component. The early effect is thought to reflect the efficiency with which words in Katakana script make contact with sublexical phonological representations involved in spoken language comprehension, due to the particular way this script is used by Japanese readers. This demonstrates fast-acting influences of visual primes on the processing of auditory target words, and suggests that briefly presented visual primes can influence sublexical processing of auditory target words. The later N400 priming effects, on the other hand, most likely reflect cross-modal influences on activity at the level of whole-word phonology and semantics.

1. Introduction

In modern societies, adult speakers of a given language are generally also fluent readers of that language. It might therefore come as no surprise that certain contemporary accounts of word-level processing during language comprehension highlight the role played by cross-modal interactions in the auditory and visual domains. According to one such account, the bi-modal interactive-activation model (BIAM: Diependaele, Ziegler, & Grainger, 2010; Grainger & Ferrand, 1994; Grainger & Holcomb, 2009; Kiyonaga, Grainger, Midgley, & Holcomb, 2007), a printed word rapidly activates associated phonological representations – both sublexical (phonemes, syllables) and lexical (whole-word phonology). Similarly a spoken word rapidly activates associated orthographic representations – both sublexical and lexical. This is in line with the known connectivity across brain regions thought to be specialized in orthographic and phonological processing (e.g., Cohen & Dehaene, 2009).

Consistent with this theorizing, there is a long history of research that has established a key role for phonological information during silent reading (e.g., Ferrand & Grainger, 1992; Perfetti & Bell, 1991; see Rastle & Brysbaert, 2006; for a

* Corresponding author. Gabrieli Lab, Brain and Cognitive Sciences, Massachusetts Institute of Technology, 43 Vassar Street 46-4033, Cambridge, MA, 02139, USA.

E-mail address: kokano@mit.edu (K. Okano).

http://dx.doi.org/10.1016/j.jneuroling.2015.09.002
0911-6044/

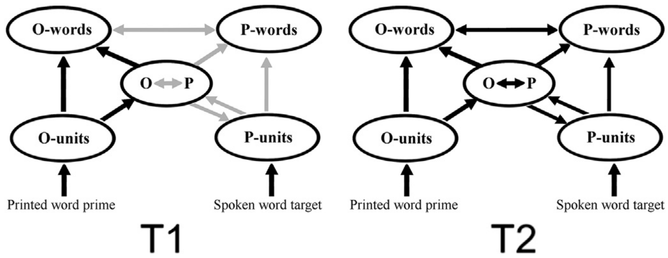

Fig. 1. Hypothetical activation levels in sublexical (P-units) and lexical (P-words) phonological representations at two moments in time (T1 & T2) during processing of an auditory target word following the brief presentation of a printed word prime. At T1, presentation of the printed word prime initiates processing of sublexical phonology associated with that word via the sublexical interface between orthography and phonology (O ⇔ P), but activation from the prime is too weak to influence processing of the auditory target word (the grey arrows signify weak activation). At T2, prime-generated activation has fully propagated to sublexical and lexical phonological representations, but processing of the spoken word is now centered on whole-word phonology and its mapping onto semantics (not shown). This illustrates the difficulty in distinguishing sublexical from lexical influences of visual primes on auditory target processing.

review). In these studies, a briefly presented pattern-masked visual prime stimulus is presented immediately prior to visual target word presentation, and the phonological overlap across prime and target is manipulated in order to investigate fast phonological influences on visual word recognition. The BIAM is the only implemented model to date that can simulate these fast phonological influences on visual word recognition as revealed by masked priming (see Diependaele et al., 2010). The model's ability to simulate these priming effects is due to the rapid sublexical conversion of spelling-to-sound, combined with the influence that this process has on activity in whole-word orthographic representations. This influence is shown in Fig. 1, which illustrates the time-course of priming effects from a visually presented prime stimulus. The left panel of Fig. 1 shows a first phase (T1) of purely orthographic influences followed by a second phase (T2) of phonological influences on both visual and auditory word recognition shown in the right panel of Fig. 1. Along with behavioral results, the combination of masked priming methodology and event-related potential (ERP) recordings has helped specify the time-course of priming effects (e.g., McPherson & Holcomb, 1999; Misra & Holcomb, 2003; see Grainger & Holcomb, 2009; for a review). More specifically, the predicted time-course of priming effects for visual word recognition (early orthographic effects followed by phonological priming) has been confirmed in several priming studies manipulating prime duration and using ERPs to track the time-course of priming effects (e.g., Grainger, Kiyonaga, & Holcomb, 2006).

One key characteristic of the BIAM is that the phonological representations activated upon presentation of a visual stimulus are the same representations as used in spoken language comprehension. It is this particularity of the BIAM that provides the focus of the present study, and that is illustrated in Fig. 1. One important point to note about Fig. 1 is that the sublexical interface between orthography and phonology (O ⇔ P) involves phonological representations that are used in spoken word recognition. This implies that as soon as a printed stimulus makes contact with sublexical phonology, it should then influence auditory word perception, and these influences should arise at the level of sublexical phonological processing as well as lexical processing. What, then, is the evidence for an early influence of visually presented primes on spoken language comprehension?

Perhaps the clearest evidence to date has been obtained using masked cross-modal priming (Grainger, Diependaele, Spinelli, Ferrand, & Farioli, 2003; Kiyonaga et al., 2007). In this paradigm, to be used in the present study (see Fig. 2), a briefly presented and pattern-masked visual prime stimulus is rapidly succeeded by an auditory target stimulus, and the influence of prime-target relatedness on target word recognition is measured. Behavioral measures of target word processing have revealed facilitatory cross-modal repetition priming with prime exposures as brief as 53 ms. Priming from pseudo-homophone primes (e.g., the nonword “brane” that can be pronounced like the real word “brain”), on the other hand, were only found to be significant with longer primes durations (67 ms). Importantly, within-modality (visual—visual) pseudo-homophone priming was also absent at the 53 ms prime duration and emerged at the longer 67 ms duration (Grainger et al., 2003).[1] This points to a minimal amount of time necessary for effective sublexical conversion of print-to-sound, and that once this conversion has taken place it influences the processing of both written and spoken target words. A later study (Kiyonaga et al., 2007) recorded ERPs to visual and auditory target words preceded by visual primes in the same masked cross-modal priming paradigm. ERPs have been shown to be differentially sensitive to both early (sublexical) and later (whole-word) processing and can therefore be used to disentangle the complex time-course of priming effects. In this study, cross-modal repetition priming effects were mostly evident on the N400 component, thought to reflect the mapping of whole-word phonological representations onto semantics (Grainger & Holcomb, 2009; Holcomb & Grainger, 2006). This suggests that cross-modal visual-auditory priming effects might arise mostly at the level of lexical representations, with little or no influence of visually presented primes on the earlier sublexical processing of auditory targets.

As illustrated in Fig. 1, according to the BIAM, sublexical influences should be observable if the sublexical translation of print-to-sound operates fast enough or is given enough time. That is, sublexical orthographic information (O-units) activated

[1] We point out here that these absolute timing values must be taken with caution, given potential differences in stimulus contrast across studies. The key message is provided by the relative timing of effects within a given study.

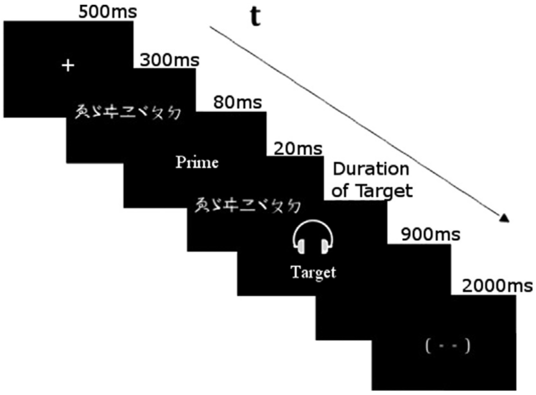

Fig. 2. Schematic of the cross-modal priming procedure used in the present study.

by the printed prime word must make contact with phonological representations involved in auditory target word processing via the sublexical interface between orthography and phonology (O ⇔ P) before such processing of the target word is complete. In the ERP study of Holcomb, Anderson, and Grainger (2005), this was achieved by using a relatively long prime presentation duration of 200 ms and having primes immediately followed by auditory targets. Repetition priming was found to emerge at around 150 ms post-target onset in these conditions. However, it is difficult to distinguish whether this effect was a result of early sublexical effects or from equally early lexical effects, such as those reported by Braun, Hutzler, Ziegler, Dambacher, and Jacobs (2009) and Hauk, Davis, Ford, Pulvermuller, and Marslen-Wilson (2006) since there was no manipulation that could point to a likely lexical or sublexical locus of the observed priming effects. The aim of the present study was therefore to seek more direct evidence for sublexical effects in cross-modal priming. This evidence will take the form of a modulation of early ERP priming effects by a factor that ought to affect the speed of sublexical conversion of print-to-sound, plus the absence of any lexical influences on the same priming effects. It is important to note that the primary focus will be the systematic examination of the time-course of the neuro-cognitive processes underlying word comprehension as opposed to identifying the neural sources of such effects (Grainger & Holcomb, 2009).

To investigate the lexical and sublexical locus of effects during cross-modal priming, we used the Japanese writing system. The Japanese writing system has three different sets of characters, Kanji, Hiragana and Katakana. Kanji are ideographs which represent whole words or parts of words. There are approximately 5000 commonly used Kanji characters and these are taught gradually between first and twelfth grade. Hiragana and Katakana are syllabary systems where each character represents one syllable (mora). There are 46 characters in both syllabary systems and these are taught in the first grade. Hiragana is used to write words for which there is no Kanji for and for word inflections, whereas Katakana is mainly used to transcribe foreign loan words and to write out onomatopoeia. Hiragana and Katakana are visually dissimilar but phonologically equivalent; therefore, it is possible to write the same word in two different scripts however, only one script would be visually familiar to experienced Japanese readers. In this study, we focus on the two Kana scripts, Hiragana and Katakana. Although native Japanese speakers are fluent in all three scripts, these scripts are thought to engage different processing mechanisms. For example, multiple clinical case studies have reported Japanese aphasic patients having different levels of reading impairments in the three scripts. The level of impairment in each script differs in each case with no consistent pattern (Endo et al., 2001; Maeshima et al., 2002; Yamadori, 2011). When focusing on the two Kana scripts, Hatta and Ogawa (1983) reported greater repetition priming in Hiragana compared to Katakana in a lexical decision task. Furthermore, Hatta (1985) reported larger Stroop interference effects in Hiragana and Kanji compared to Katakana. Taken together, one reasonable conclusion from these studies is that Hiragana may engage more lexical processing than Katakana.

In the present study, masked prime stimuli were Japanese words written in one of the two syllabary scripts of Japanese (Hiragana and Katakana). Because there has not been much psycholinguistic research contrasting Hiragana and Katakana (and even less using cognitive neuroscience techniques) the hypothesis of this study was mainly guided by the general preferences of use of the two scripts. Fifty percent of prime words in each script were items typically written in that script and 50 percent were items not typically written in that script. Targets were spoken Japanese words to which ERPs were recorded. Use of Katakana script is particularly important here, since we expect sublexical translation of orthography to phonology to be highly optimized in this script, given that compared with both Hiragana and Kanji scripts, Katakana is mostly used to generate a phonological representation from print (e.g., it is used to write onomatopoeias and transcribe foreign loan words (Shinmura, 2008)). This is therefore expected to increase the possibility of observing early cross-modal priming effects during spoken word recognition, as predicted by the BIAM. Differences in priming effects between Hiragana and Katakana primes, as well as effects of prime word familiarity (i.e., whether or not the spoken word is typically written in that script or not), will provide further evidence with respect to the lexical or sublexical locus of effects. As it is depicted in Fig. 1, it is extremely difficult to disentangle sublexical versus lexical influences of visual primes in auditory target processing. Japanese Kana provides a platform where a single word can be written in two different scripts, only one of which is the preferred script for a given word. If it is the case that Katakana has stronger sublexical associations between orthography and phonology (O ⇔ P connection in

Fig. 1) than Hiragana, due to the way this script is primarily used, then we expect to see an early influence of Katakana primes on auditory target processing, an influence that should be earlier than observed in prior research with English stimuli (Kiyonaga et al., 2007) and earlier than the expected effects of Hiragana primes in the present study. Furthermore, our manipulation of script typicality will allow us to evaluate the influence of whole-word orthographic familiarity on priming effects, since primes presented in the script in which they are typically read are expected to involve more direct access to whole-word orthographic representations. Primes written in the non-typical script, on the other hand, are more likely to be read via sublexical conversion of orthography to phonology.

2. Methods and materials

2.1. Participants

Twenty-one native Japanese speakers currently residing in the greater Boston, MA area were recruited and paid for their participation (Age M = 27.19, SD = 4.88, Male = 9). All participants were right-handed with normal or corrected-to-normal vision, reporting Japanese as their native language, with no history of language or reading disorders. All tests and instruments were reviewed and approved by the Tufts Internal Review Board. Informed consent was obtained from each participant prior to the experiment.

2.2. Stimuli

The critical stimuli for this experiment were derived from 480 Japanese words (Okano, Grainger, & Holcomb, 2013). Visual primes were 120 distinctly Hiragana words presented in Hiragana (Hh), 120 distinctly Katakana words presented in Katakana (Kk), 120 typically Hiragana words presented in Katakana (Hk), and 120 typically Katakana words presented in Hiragana (Kh). These visual primes were paired with auditory target words that were either repeats of the prime or were unrelated to the prime. The unrelated visual primes were phonologically, orthographically, morphologically, and semantically unrelated to the paired auditory target word. There were a total of 480 pairs of repeated words and 480 pairs of unrelated words. For purposes of counterbalancing, each participant was presented with one of three lists. All lists contained all eight conditions (Hh repeated, Hh unrelated, Kk repeated, Kk unrelated, Hk repeated, Hk unrelated, Kh repeated, Kh unrelated) with each prime word appearing only once. The auditory target words were digitized speech (22 KHz) recorded by a female native Japanese speaker with a mean duration of 533 ms (range 277–783 ms) for Hiragana words and 534 ms (range 300–702) for Katakana words.

2.3. Task

Participants completed a go/no-go semantic categorization, where they were asked to press a button as soon as they encountered a Japanese word corresponding to a body part. For all other words (i.e., the critical items described above), participants were asked to listen/view passively without responding. Fifteen percent of the total number of auditory targets and 10% of the total number of visual primes were body part probe words. A brief practice session preceded the experiment to familiarize the participant with the task.

2.4. Procedure

Visual stimuli were presented on a 19-in CRT monitor with a refresh rate of 100 Hz and located 56 inches in front of the participant. Stimuli were displayed in off-white letters in Mincho font (letter matrix 20 pixels wide $\times$ 40 pixels tall) on a black background. Each trial began with a fixation cross displayed in the center of the screen for 500 ms. Then a forward mask comprised of eight pseudo characters replaced the fixation cross for 300 ms.

The forward mask was then replaced by a prime for 80 ms, which was quickly replaced by the backward mask (same string of pseudo characters as the forward mask) for 20 ms. The backward mask was then replaced by a blank screen at which time the subject heard an auditory target word. The auditory target word was followed by a 900 ms blank screen which was replaced by a blink sign (- -) for 2000 ms informing participants that they were now allowed to blink. The procedure is summarized in Fig. 2.

2.5. EEG recording procedure

Participants were seated in a comfortable chair in a sound attenuated darkened room. The electroencephalogram (EEG) was recorded from 29 electrodes held in place on the scalp by an elastic cap (Electro-Cap International, Inc., Eaton, OH, USA) in the standard 10–20 layout. In addition to the 29 scalp sites, additional electrodes were attached to below the left eye (to monitor for vertical eye movement/blinks), to the right of the right eye (to monitor for horizontal eye movements), over the left mastoid bone (reference) and over the right mastoid bone (recorded actively to monitor for differential mastoid activity). All EEG electrode impedances were maintained below 5 kΩ (impedance for eye electrodes was less than 10 kΩ). The EEG was amplified by an SA Bioamplifier (San Diego, CA, USA) with a bandpass of 0.01 and 40 Hz, and the EEG was continuously sampled at a rate of 200 Hz throughout the experiment.

2.6. Data analysis

Averaged ERPs time-locked to target onset were formed off-line from trials free of ocular/muscular artifact and were lowpass filtered at 15 Hz, and a 100 ms epoch immediately before target onset was used as the baseline. An average of 89.2% (SD = 9.7%) of the trials per participant were used in the analysis and were evenly distributed across conditions. We analyzed mean amplitudes in three contiguous temporal epochs following target word onset that in a number of previous studies have been demonstrated to capture priming effects (e.g., Grainger et al., 2006): 180–210 ms, 210–300 ms and 300–400 ms. To make the analysis of electrode location manageable we used an ROI approach, which involved averaging across four electrode sites at each of four scalp regions (anterior, posterior, left lateral, and right lateral – see Fig. 3 caption). Repeated measures ANOVAs started with four factors: ROI (anterior vs. posterior vs. left lateral vs. right lateral), Prime Script (Hiragana vs. Katakana), Script Typicality (typical vs. atypical), and Repetition (repeated vs. unrelated). In the case of significant interactions involving the Repetition, Prime Script, Script Typicality or ROI factors, additional planned follow-up ANOVAs were performed. The Geisser-Greenhouse correction was applied to all contrasts involving more than 1 degree of freedom in the numerator (Greenhouse & Geisser, 1959). For a more complete picture of the scalp distribution of priming effects, we also plotted the full montage of sites as interpolated voltage maps.

3. Results

3.1. Behavioral data

Participants detected 92% (SD = 7%) of human body part probes in the auditory target words. All subjects reported seeing prime stimuli, and 16 subjects detected body part probe words among the visual primes (M = 48%, SD = 25%). The detection

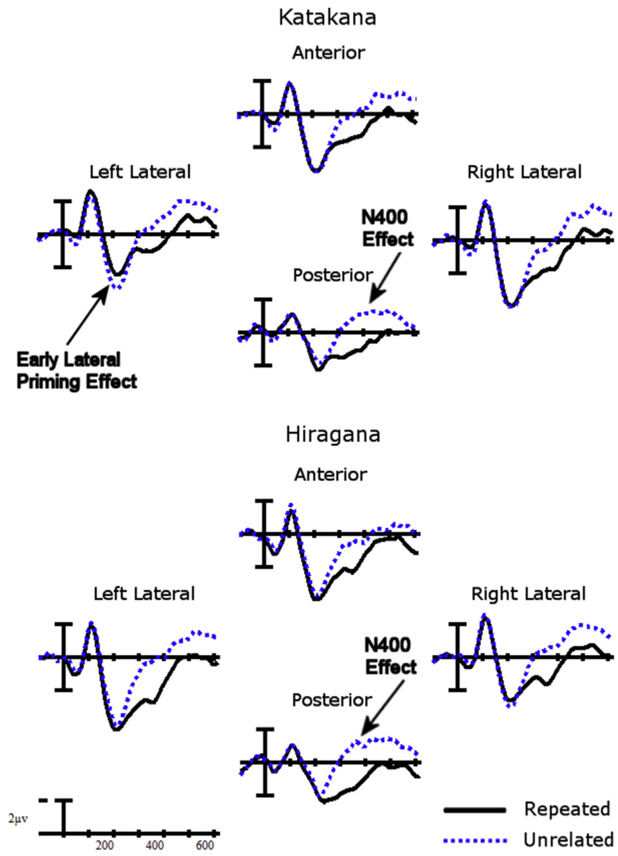

Fig. 3. ERPs from all repeated (solid black line) and unrelated (dotted blue line) targets in the Katakana condition (top panel) and the Hiragana condition (bottom panel) at 4 ROIs. Each ROI included 4 electrodes: Anterior ROI included electrode sites FPz, FP1, FP2, FZ, Posterior ROI included sites Pz, O1, O2, Oz, Left Lateral ROI included sites FC5, T3, C3, CP5, and Right Lateral ROI included sites FC6, C4, T4, CP6. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

rates of these probe words in Hiragana (M = 2.56, SD = 1.36) and Katakana (M = 3.19, SD = 1.94) were not significantly different $t(15) = 1.50$, $p = .15$.

3.2. Visual inspection of ERPs

Plotted in Fig. 3 are the ERPs for the repeated and unrelated conditions in the four ROIs in Hiragana and Katakana conditions respectively. Plotted in Fig. 4 are blow-ups of the FC5 scalp site, found to have the peak amplitude during the time window of interest, and accompanying scalp voltage map. As can be seen, there is a clear positive deflection around 200 ms that was modulated by repetition for Katakana primes. Fig. 5 shows a negative deflection starting around 200 ms post-target onset (N400) that was influenced by repetition for both Hiragana and Katakana primes.

3.3. Analyses of ERP data

3.3.1. 180–210 ms epoch

The ANOVA in this epoch revealed a significant Prime Script × Repetition × ROI interaction ($F(3,60) = 2.79$, $p = .047$). Examination of the ERPs in Fig. 3 suggests that repetition priming effects were largest in the left lateral ROI in the Katakana prime script condition. Consistent with this observation, follow-up analyses looking only at the left lateral ROI indicated that the Repetition effect for auditory targets following Katakana primes ($F(1,20) = 6.38$, $p < .02$) was stronger than for Hiragana primes ($F(1,20) = 1.13$, $p > .3$). All other effects in both the Katakana and Hiragana prime conditions were not significant (all $Fs(1,20) < 2.47$, $p > .13$).

3.3.2. 210–300 ms epoch

Between 210 and 300 ms, we found a significant main effect of Repetition $F = 4.43$ (1, 20) $p < .05$, indicating that targets following unrelated primes were more negative going than targets that were repetitions of primes. We also found a Repetition × ROI interaction, $F = 5.62$ (3, 60) $p < .005$. In order to decompose this interaction we ran follow-up ANOVAs which revealed that the effects of Repetition tended to be larger over posterior s than anterior ($F = 5.52$(1, 20) $p < .05$) or left lateral ($F = 12.57$ (1, 20) $p < .005$) sites, and larger over right lateral than left lateral sites ($F = 6.06$ (1, 20) $p < .05$ – all other effects were not significant, all $Fs < 2.3$, $p > .1$).

3.3.3. 300–500 ms (N400) epoch

Between 300 and 500 ms, we found a significant main effect of Repetition $F = 23.33$ (1, 20) $p < .001$, indicating that targets following unrelated primes were more negative going than targets that were repetitions of primes. We also found a Repetition × ROI interaction, $F = 4.78$ (3, 60) $p < .005$. Follow-up analyses revealed that the effects of Repetition were larger over posterior than anterior sites ($F = 8.62$(1, 20) $p < .01$); larger over posterior than left lateral sites ($F = 4.63$ (1, 20) $p < .05$) and larger over right lateral than anterior sites ($F = 6.54$ (1, 20) $p < .05$; all other contrasts were not significant – all $Fs < 2.2$, $p > .1$).

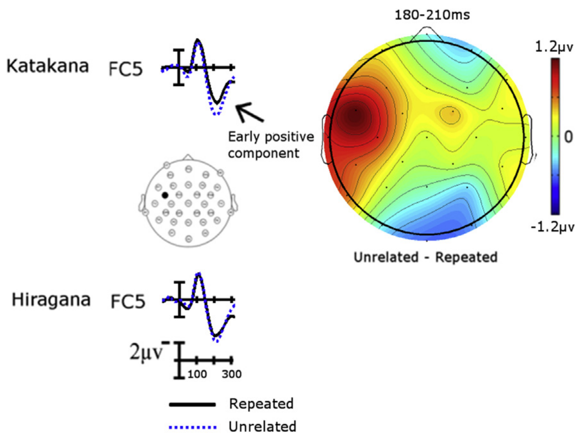

Fig. 4. ERP cross-modal repetition priming effects for the Katakana primes (upper panel) and Hiragana primes (lower panel). The voltage map on the right is from all scalp sites based on average activity of the unrelated prime condition subtracted from the repeated prime condition averaged across 180–210 ms in the Katakana condition. ERPs plotted on the left are from the repeated (solid black line) and unrelated (dotted blue line) prime conditions at the FC5 site for Katakana and Hiragana. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

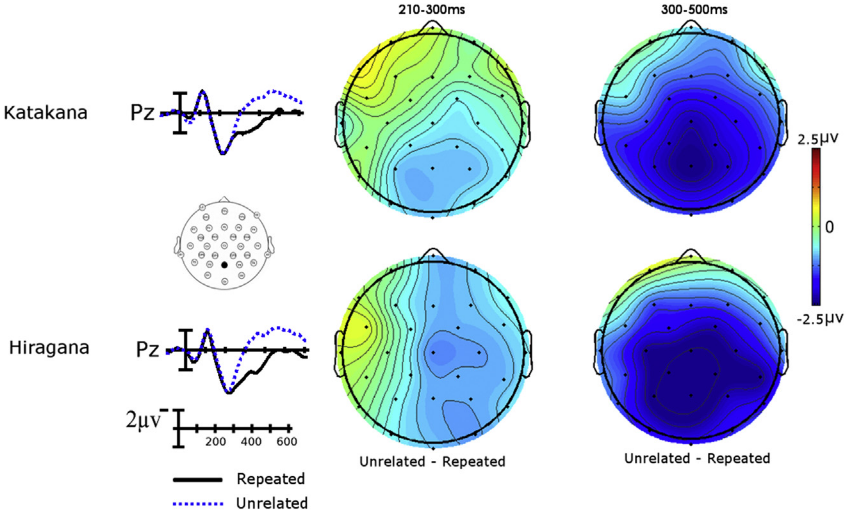

Fig. 5. ERP Repetition effects in two epochs 210–300 ms and 300–500 ms (N400) component for all Katakana prime (upper panels) and Hiragana prime (lower panels) condition. ERPs plotted on the left are from repeated (solid black line) and unrelated (dotted blue line) targets at the Pz site. The voltage maps on the right are from all scalp sites based on average activity of unrelated minus repeated target ERPs averaged across 300–500 ms. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

4. Discussion

In the present study ERPs were recorded while native Japanese participants listened to Japanese words preceded by masked visual primes. These prime words could either be the same word as the auditory target or a different word, and could be written in Hiragana or Katakana script. Repetition priming effects were seen in ERP waveforms, with an initial focal effect peaking at around 200 ms post-target onset, followed by a temporally and spatially more widespread priming effects on the N400 component. The pattern seen on the N400 component replicates prior work on masked cross-modal priming in French (Kiyonaga et al., 2007), and likely reflects the influence of visual primes on lexical processing of auditory targets. The fact that cross-modal priming effects observed in the present study were not influenced by prime script typicality (i.e., whether the word is typically written in that particular script or not), suggests that the transfer of information across modalities was mostly operating sublexically. Our measure of prime script typicality reflects the relative frequency with which a given Japanese word is written in a given script. If the priming effects were being driven mainly by lexical processing, then we should have seen greater priming effects for primes presented in the script in which they are typically read. Processing of primes written in the typical script will rely more on direct access to whole-word orthographic representations (see Fig. 1) than primes written in the atypical script, and this will provide an additional source of input to whole-word phonological representations that can combine with information transferred from the sublexical interface between orthography and phonology. The absence of an interaction between prime script typicality and priming therefore suggests that the bulk of the priming effects were being driven by the sublexical conversion of orthography-to-phonology (the O ⇔ P interface in Fig. 1) influencing auditory target word processing. Primes presented in typical and atypical format would be processed equally well along this pathway. These conclusions must nevertheless be taken with caution given that they are based on a null effect.

A key question that was addressed in the present research was whether or not the sublexical conversion of orthography-to-phonology could operate rapidly enough in order to influence sublexical processing of auditory target words. The use of primes written in Katakana script was expected to optimize processing along the sublexical interface between orthography and phonology, and therefore to provide an opportunity to observe earlier cross-modal priming effects than previously reported in English (Kiyonaga et al., 2007), where effects were only seen on the N400 component. Kiyonaga et al. (2007) interpreted these cross-modal N400 priming effects as reflecting influences operating at the level of whole-word phonological representations, in line with other reports of modulation of N400 amplitude by lexical manipulations in single word reading paradigms (e.g., Holcomb, O'Rourke, & Grainger, 2002; Laszlo & Federmeier, 2011). The early effect of repetition priming found with Katakana primes in the present study therefore provides the first clear evidence for cross-modal transfer influencing the sublexical processing of auditorily presented target words.

The fact that the early priming effect was only evident with primes printed in Katakana script is likely due to the fact that this particular script is used specifically for "sounding out" unfamiliar words (e.g., transcribing foreign loan words and writing onomatopoeias). This implies that the efficiency of the sublexical mapping of print-to-sound (O-units → O ⇔ P interface → P-words) may be greater in this script. It is in these precise conditions that we expected to observe early cross-modal priming effects, given the architecture of the BIAM (see Fig. 1). As indicated in Fig. 3, the early priming effect is of the opposite polarity to the N400 priming effects. This, in addition to the clearly distinct spatial distribution of the effect, points to different underlying neural generators compared with the lexically driven N400 effects. Furthermore, the different pattern of early priming effects seen for Katakana and Hiragana in the present study contrasts with the absence of any differential priming for the two scripts reported in our prior work with visual primes and targets (Okano et al., 2013). Most important is that cross-script repetition priming effects (i.e., the same spoken word written in one script for the prime and the other script for the target) were found not to interact with script, and emerged later (in the N250 time window) than the early priming effect of the present study. Okano et al. (2013) argued that the cross-script N250 priming effects they observed were likely driven by the phonology shared by primes and targets in different scripts. This therefore suggests that the early priming effect found in the present study is likely subtended by a different mechanism. One possibility is that Katakana script is strongly connected with articulatory representations that in turn are connected with sublexical phonological representations involved in spoken word recognition. Direct connectivity between sublexical orthography and articulatory output has been proposed in order to explain masked onset priming (Grainger & Ferrand, 1996), a phenomenon whereby primes that share their initial phoneme with target words (e.g., torp-table) facilitate target processing relative to primes with a different initial phoneme (morp-table), but only when the task requires articulatory output (Carreiras, Ferrand, Grainger, & Perea, 2005; Kinoshita & Woolams, 2002; Schiller, 2008). Masked onset priming can be obtained with very brief prime durations, and is known to arise very rapidly (Klein et al., 2015). This account predicts that Katakana primes should reveal masked onset priming effects that are just as strong in silent reading and naming, whereas Hiragana primes should show the typically stronger onset priming effects in naming.

One problem, however, for a purely articulatory account of the early Katakana priming effect, is its focal nature, localized in left temporal electrode sites. This suggests that the effect might be driven by neural structures located in or near left superior temporal or middle temporal cortex. While it has been suggested that these regions might be involved in lexico-semantic processing (e.g., Indefrey & Levelt, 2004; Gold, Balota, Kirchhoff, & Buckner, 2005), given the evidence reported here that the early priming effect is distinct from the N400 lexical effect, we think our result is most consistent with the findings of previous MEG and fMRI studies suggesting that activity in this region is involved in phonological processing (e.g., Gow, 2012; Hickok & Poeppel, 2007). Within the architecture for cross-modal interactions described in Fig. 1, this early priming effect is located at the level of word-centered sublexical phonological representations that mediate between incoming acoustic input and whole-word phonology during spoken word recognition. Such representations would code for the position of phonemes in words, while abstracting away from the precise timing of the auditory input. It is these representations that can be mapped onto the equivalent word-centered sublexical orthographic code via the sublexical O ⇔ P interface in the BIAM. Here it is important to note that a new model of spoken word recognition, incorporating this kind of sublexical word-centered phonological code, has recently been developed and successfully tested against the standard benchmarks in this field (Hannagan, Magnuson, & Grainger, 2013). In sum, within the framework of the BIAM (Fig. 1), the early priming effect would be driven by activation in word-centered sublexical phonological representations (P-units), and processing of Katakana print would be optimized for this. N400 priming effects, on the other hand, would reflect transmission of information between the sublexical interface between orthography and phonology (O ⇔ P) and whole-word phonological representations (P-words), and processing of Hiragana print would be optimized for this.

It is also interesting to compare the early priming effect seen with Katakana primes in the present study with prior reports of early within-modality repetition priming effects in the visual and auditory modalities. The timing, direction, and topography of the present effect (see Fig. 4) clearly distinguishes it from N/P150 and N250 priming effects found with visual masked primes and visual targets (e.g., Chauncey, Holcomb, & Grainger, 2008; see Grainger & Holcomb, 2009; for review). Furthermore, although the timing is similar, the direction and topography of the effect distinguishes it from the earliest repetition priming effects found with auditory primes and targets using a dichotic listening procedure (Grainger & Holcomb, 2015). Therefore, although the timing of the early cross-modal priming effect points to sublexical phonological processing as a likely source, it is clear that the manner in which such sublexical phonology is activated by prime stimuli (via orthographic representations or via acoustic representations) greatly determines the way such processing manifests itself in the ERP waveforms.

More generally speaking, the present findings are in line with the growing literature demonstrating orthographic influences on spoken word recognition in the absence of any visual stimulus whatsoever. Early work demonstrated that it is harder to judge that two spoken words rhyme when there is a mismatch in orthography (e.g., Seidenberg & Tanenhaus, 1979). Ziegler and Ferrand (1998) demonstrated that the time to recognize auditorily presented words is affected by the consistency with which the word's phonemes map onto orthographic representations (e.g., the phonological rhyme of the French word “dos” has several different orthographic realizations in words such as “peau”, “saut”, “maux”). It has also been shown that the number of orthographic neighbors of spoken word targets affects lexical decisions to these words independently of their phonological neighborhood density (Ziegler, Muneaux, & Grainger, 2003; Grainger, Muneaux, Farioli, & Ziegler, 2005). Furthermore, auditory lexical decisions are influenced by the degree of orthographic overlap of auditorily presented primes (Chéreau, Gaskell, & Dumay, 2007; Jakimik, Cole, & Rudnicky, 1985; Perre, Midgley, & Ziegler, 2009; Slowiaczek, Soltano, Wieting, & Bishop, 2003), and such orthographic influences can be obtained in conditions where participants are not aware of the orthographic relationship between primes and targets (Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008).

Finally, ERP studies of orthographic influences on spoken word recognition have revealed how the effects of orthographic inconsistency emerge on-line during the processing of auditorily presented target words (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Perre & Ziegler, 2008). In these studies, the earliest orthographic effects emerged at around 300 ms post-target onset, and depended on the location of the inconsistent information in the spoken word. Overall, the effects were present at about 200 ms after the arrival of the inconsistent orthographic information in the auditory signal. Within the framework of the BIAM, these effects would start to arise at the sublexical O ⇔ P interface and would subsequently reflect the mapping of sublexical phonology onto whole-word phonological representations (N400 effects). The priming methodology used in the present study, combined with the characteristics of Katakana words, enabled us to detect even earlier orthographic influences on spoken word recognition.

Acknowledgment

This research was supported by grant number HD25889 awarded to PJH. JG was supported by ERC grant 230313.

References

Braun, B., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping, 30*, 1977–1989.
Carreiras, M., Ferrand, L., Grainger, J., & Perea, M. (2005). Sequential effects of phonological priming in visual word recognition. *Psychological Science, 16*, 585–589.
Chauncey, K., Holcomb, P. J., & Grainger, J. (2008). Effects of stimulus font and size on masked repetition priming: an ERP investigation. *Language and Cognitive Processes, 23*, 183–200.
Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: orthographic effects in auditory priming. *Cognition, 102*, 341–360.
Cohen, L., & Dehaene, S. (2009). Ventral and dorsal contributions to word reading. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 789–804). Cambridge, MA: MIT Press.
Diependaele, K., Ziegler, J., & Grainger, J. (2010). Fast Phonology and the Bi-modal Interactive Activation Model. *European Journal of Cognitive Psychology, 22*, 764–778.
Endo, K., Suzuki, K., Yamadori, A., Kumabe, T., Seki, K., & Fujii, T. (2001). A case of crossed aphasia with echolalia after the resection of tumor in the right medial frontal lobe. *No To Shinkei, 53*, 287–292.
Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology, 45A*, 353–372.
Gold, B. T., Balota, D. A., Kirchhoff, B. A., & Buckner, R. L. (2005). Common and dissociable activation pattern associated with controlled semantic and phonological processing: evidence from FMRI adaptation. *Cerebral Cortex, 15*, 1438–1450.
Gow, D. W., Jr. (2012). The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. *Brain Language, 121*, 273–288.
Grainger, J., Diependaele, K., Spinelli, E., Ferrand, L., & Farioli, F. (2003). Masked repetition and phonological priming within and across modalities. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 1256–1269.
Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: effects of masked homophone primes. *Journal of Memory & Language, 33*, 218–233.
Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: cross-task comparisons. *Journal of Memory and Language, 35*, 623–647.
Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: on the time-course of component processes in visual word recognition. *Language and Linguistics Compass, 3*, 128–156.
Grainger, J., & Holcomb, P. J. (2015). An ERP investigation of dichotic repetition priming with temporally overlapping stimuli. *Psychonomic Bulletin and Review, 22*, 289–296.
Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time-course of orthographic and phonological code activation. *Psychological Science, 17*, 1021–1026.
Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *Quarterly Journal of Experimental Psychology, 58A*, 981–998.
Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.
Hannagan, T., Magnuson, J., & Grainger, J. (2013). Spoken word recognition without a TRACE. *Frontiers in Language Sciences, 4*, 563.
Hatta, T. (1985). Reading processes in Japanese: do the Japanese have script-specific mechanisms? *Language Sciences, 7*, 355–363.
Hatta, T., & Ogawa, T. (1983). Hiragana and Katakana in Japanese orthography and lexical representation. *Language Sciences, 5*, 185–196.
Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word-recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*, 1383–1400.
Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Review Neuroscience, 8*, 393–402.
Holcomb, P. J., Anderson, J., & Grainger, J. (2005). An electrophysiological investigation of cross-modal repetition priming. *Psychophysiology, 42*, 493–507.
Holcomb, P. J., & Grainger, J. (2006). On the time-course of visual word recognition: an ERP investigation using masked repetition priming. *Journal of Cognitive Neuroscience, 18*, 1631–1643.
Holcomb, P. J., O'Rourke, T., & Grainger, J. (2002). An event-related brain potential study of orthographic similarity. *Journal of Cognitive Neuroscience, 14*, 938–950.
Indefrey, O., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.
Jakimik, J., Cole, R. A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language, 24*, 165–178.
Kinoshita, S., & Woolams, A. (2002). The masked onset priming effect in naming: computation of phonology or speech planning? *Memory and Cognition, 30*, 237–245.
Kiyonaga, K., Grainger, J., Midgley, K. J., & Holcomb, P. J. (2007). Masked cross-modal repetition priming: an event-related potential investigation. *Language and Cognitive Processes, 22*, 337–376.
Klein, M., Grainger, J., Wheat, K. L., Millman, R. E., Simpson, M. I. G., Hansen, P. C., et al. (2015). Early activity in Broca's area during reading reflects fast access to articulatory codes from print. *Cerebral Cortex, 25*, 1715–1723.
Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology, 48*, 176–186.
Maeshima, S., Toshiro, H., Sekiguchi, E., Okita, R., Yamaga, H., Ozaki, F., et al. (2002). Transcortical mixed aphasia due to cerebral infarction in left inferior frontal love and temporo-parietal lobe. *Neuroadiology, 44*, 133–137.
McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology, 36*, 53–65.
Misra, M., & Holcomb, P. J. (2003). Event-related potential indices of masked repetition priming. *Psychophysiology, 40*, 115–130.
Okano, K., Grainger, J., & Holcomb, P. J. (2013). An ERP investigation of visual word recognition in syllabary scripts. *Cognitive Affective and Behavioral Neuroscience, 13*, 390–404.
Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience, 21*, 169–179.
Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: evidence from backward masking and priming. *Journal of Memory & Language, 30*, 473–485.

Perre, L., Midgley, K., & Ziegler, J. C. (2009). When beef primes reef more than leaf: orthographic information affects phonological priming in spoken word recognition. *Psychophysiology, 46*, 739–746.

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research, 1188*, 132–138.

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

Schiller, N. O. (2008). The masked onset priming effect in picture naming. *Cognition, 106*, 952–962.

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 546–554.

Shinmura, I. (2008). *Kojien* (6th ed.). Tokyo: Iwanami Shoten.

Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of orthography and phonology in spoken word recognition. *Quarterly Journal of Experimental Psychology, 56A*, 233–262.

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: pseudohomograph priming. *Journal of Memory and Language, 58*, 366–379.

Yamadori, A. (2011). Gogi (word-meaning) aphasia. *Brain Nerve, 63*, 811–820.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: the consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*, 683–689.

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