CrossMark

Role of consciousness in temporal integration of semantic information

Yung-Hao Yang[1] · Yung-Hsuan Tien[1] · Pei-Ling Yang[1] · Su-Ling Yeh[1,2,3]

Abstract Previous studies found that word meaning can be processed unconsciously. Yet it remains unknown whether temporally segregated words can be integrated into a holistic meaningful phrase without consciousness. The first four experiments were designed to examine this by sequentially presenting the first three words of Chinese four-word idioms as prime to one eye and dynamic Mondrians to the other (i.e., the continuous flash suppression paradigm; CFS). An unmasked target word followed the three masked words in a lexical decision task. Results from such invisible (CFS) condition were compared with the visible condition where the preceding words were superimposed on the Mondrians and presented to both eyes. Lower performance in behavioral experiments and larger N400 event-related potentials (ERP) component for incongruent- than congruent-ending words were found in the visible condition. However, no such congruency effect was found in the invisible condition, even with enhanced statistical power and top-down attention, and with several potential confounding factors (contrast-dependent processing, long interval, no conscious training) excluded. Experiment 5 demonstrated that familiarity of word orientation without temporal integration can be processed unconsciously, excluding the possibility of general insensitivity of our paradigm. The overall result pattern therefore suggests that consciousness plays an important role in semantic temporal integration in the conditions we tested.

While we tend to fixate on one word at a time during reading, meanings of parafoveal words that are not consciously perceived are still extracted and could affect the reading process. Specifically, studies on semantic preview benefits (Hohenstein & Kliegl, 2014; Rayner & Schotter, 2014; Yan, Richter, Shu, & Kliegl, 2009; Yan, Zhou, Shu, & Kliegl, 2012) demonstrated that semantic information of the preceding parafoveal words actually provided clues about the upcoming target words as indexed by shorter fixation durations when parafoveal words were semantically related to target words than when they were unrelated. A study focusing on visual crowding (Yeh, He, & Cavanagh, 2012) also illustrated that a crowded prime word still robustly yielded semantic priming effect on the subsequent target even though the prime word was unrecognizable when presented in the peripheral visual field and surrounded by flankers. These findings suggest that meanings of words are not completely blocked even when words per se are not consciously identified.

However, it remains unknown whether meanings of temporally segregated words can also be unconsciously integrated into meaningful phrases. Since temporal integration is important in language processing, the current study aims to take an in-depth examination on the possibility of unconscious integration of temporally segregated words. Regarding this issue, recent studies on unconscious semantic integration using

✉ Su-Ling Yeh
  suling@ntu.edu.tw

[1] Department of Psychology, National Taiwan University, Taipei, Taiwan

[2] Graduate Institute of Brain and Mind Sciences, National Taiwan University, Taipei, Taiwan

[3] Neurobiology and Cognitive Science Center, National Taiwan University, Taipei, Taiwan

sequentially presented stimuli in the backward-masking paradigm yielded divergent results. In one study, Kawakami and Yoshida (2015) showed that sequential picture frames could be unconsciously integrated into a socially meaningful story, and the characters' attributes could also be automatically inferred. In their experiment, participants were presented with a series of subliminal pictures that represented a story of one-character chasing after or beating the other one in a tug-of-war. Although the participants did not perceive any pictures, they tended to associate each character appropriately with its specific characteristics (e.g., chaser as more active and powerful) in the Implicit Association Test. However, the ability to extract social meanings behind subliminal picture series may rely on the accumulation of lower level processing instead of directly integrating distinct semantic information. Since the successive subliminal pictures in the study of Kawakami and Yoshida (2015) share common low-level features (Faivre & Koch, 2014b), the continuous activation among V1 may gradually build up into motion processing (Faivre & Koch, 2014a). The automatic mental associations of characters' attributes from these pictures may be inferred from the relative spatial (Kiesel, Kunde, Pohl, Berner, & Hoffmann, 2009) and temporal (Güldenpenning, Koester, Kunde, Weigelt, & Schack, 2011) relations between successive pictures.

In another study, van Gaal et al. (2014) manipulated grammatical negations (i.e., using negative modifiers, such as *not* to flip the valence of the adjective) and found that while a subliminally presented negative modifier and adjective could be integrated and generate the negation moderation effect, the effect was observed only when words were simultaneously presented (Experiment 2), but not when they were sequentially presented (Experiment 1). Even though van Gaal et al. (2014) did not observe an unconscious temporal integration of words, whether the absence of that effect was instead due to limitations from a short presentation duration and behavioral measurement is unclear. With the backward-masking paradigm, words had to be presented very briefly (e.g., 50 ms in their Experiment 1) to be invisible; such duration might be too short to allow for subliminal semantic information of each word to register and become temporally integrated. Furthermore, the failure to integrate a temporally segregated modifier (i.e., *not*) with an adjective (their Experiment 1) could also be due to the insensitivity of the behavioral responses to reflect the semantic integration processing, since van Gaal et al. (2014) did not apply electrophysiological recordings in that experiment as they did in the one with simultaneously presented words (their Experiment 2).

To render visual stimuli invisible for a relatively long duration, to allow semantic integration to occur, we used the continuous flash suppression (CFS; Tsuchiya & Koch, 2005) paradigm. In this paradigm, critical visual information from one eye is interocularly suppressed by dynamically flashing high-contrast masks from the other eye. Despite some studies showing the absence of semantic processing (e.g., Heyman & Moors, 2014; Kang, Blake, & Woodman, 2011), several recent studies adopting this paradigm have found that word frequency (Prioli & Kahan, 2015), semantic meaning (Costello, Jiang, Baartman, McGlennen, & He, 2009; Lin & Yeh, 2015; Yang & Yeh, 2011), and spatial integration of words (Sklar et al., 2012) can be processed under CFS. We reasoned that longer stimulus duration under CFS would trigger stronger and more sustained brain activation that might allow for semantic information to be integrated over time.

We chose Chinese four-word idioms as our experimental stimuli. These idioms have several properties that make them suitable for our investigation: First, all idioms are constructed with fixed four-word arrangements and cannot be understood without integrating individual words. Second, because idioms are commonly used in daily conversations and are well-known to Taiwanese, incorrect idioms can be easily identified if any of the words is substituted. Third, most of the Chinese idioms are derived from meaningful context of a historical event, an ancient legend, or a poem; hence, the holistic meaning of the idioms is more than the sum of the individual semantic information. Finally, it has been widely demonstrated that correct Chinese idioms lead to faster reaction time (called RT hereafter; Liu, Li, Shu, Zhang, & Chen, 2010; Zhang, Yang, Gu, & Ji, 2013) and smaller N400 amplitude (Liu et al., 2010; Zhou, Zhou, & Chen, 2004) than matched incorrect idioms, and the processing advantages were still preserved when the idioms were translated into a second language (Carrol & Conklin, 2014, 2015). While most of the idioms used in the current study were traditional Chinese four-character idiomatic expressions, a small subset contained commonly used Chinese four-word phrases that share similar properties (i.e., fixed arrangement and high familiarity) as the idioms.

Using CFS and Chinese four-word idioms, we examined whether sequentially presented Chinese words could be integrated into holistically meaningful idioms without consciousness using behavioral measures and event-related potentials (ERPs). In Experiments 1–4, Chinese four-word idioms were adopted with the first three words as prime and with either the original ending (the congruent condition), a semantically unrelated word (the incongruent condition), or a nonword as a target in a lexical decision task (LDT). If the meaning of the first three words could be integrated into idiom contexts, we would expect to find lower performance (slower RTs and lower accuracies) in the behavioral experiments and larger N400 amplitudes in the ERP experiments following the incongruent endings compared to the congruent endings (i.e., the congruency effect). In Experiment 1, we measured and compared both ERPs and behavioral responses to both the congruent and incongruent endings. In Experiment 2, we further tested whether top-down attention can enhance the integration of sequentially presented Chinese words. In Experiment 3, we

excluded possible confounding factors from long prime-target interstimulus interval (ISI) and contrast-dependent visibility with ERPs. In Experiment 4, we tested whether unconscious processing of semantic integration can be facilitated by preceding invisible behavioral trials with visible trials as a form of conscious training. And in Experiment 5, we presented all four words from each idiom simultaneously per trial to provide a basis for assessing the sensitivity of the paradigm used in this study.

Experiment 1

We investigated whether temporal integration of semantic information requires consciousness with ERP (Experiment 1a) and behavioral (Experiment 1a and 1b) measurements. The first three words of the Chinese four-word idioms were either rendered invisible by CFS (the invisible condition) or remained visible as superimpositions on the masks (the visible condition). The fourth word (the target of the LDT task) was unmasked, which would be either a correct ending of the idiom (the congruent trials), an incorrect ending (the incongruent trials), or a nonword. Since the idiom context of the first three words was irrelevant to the lexical judgment of the last word, the use of LDT—as an orthogonal manipulation—would be able to eliminate participant's bias (i.e., guessing the experiment's main purposes and adopting unwanted strategies).

In Experiment 1a, ERPs were recorded and time-locked to the target onset while behavioral measures were conducted at the same time. We focused on the ERP N400 component, which is a well-established index of semantic integration (see Kutas & Federmeier, 2011, for a review), since N400 amplitude is larger in cases of semantic violations. If sequentially presented words could be integrated into idiom context without consciousness, we would observe the congruency effect—larger N400s and lower performance in the incongruent trials than that in the congruent trials—in the invisible condition. However, if consciousness is necessary for word meaning integration, we would find the congruency effect only in the visible condition. Because results of the ERPs indicated clear congruency effect in the visible condition in Experiment 1a, but the behavioral measurements thereof were not as clear, we focused on behavioral performance of the invisible condition and doubled the number of trials in Experiment 1b to test whether increasing statistical power can enhance the detection of the priming effect at the behavior level.

Method

Participants Thirty healthy, naïve volunteers (18–29 years) took part in Experiment 1a, and 30 healthy naïve volunteers (18–32 years) took part in Experiment 1b. In this study, all participants were native Chinese speakers with normal or corrected-to-normal vision. They all gave informed consent before their participation. All experiments were approved by the Research Ethics Committee at the National Taiwan University (NTU REC: 201408HM005).

Stimuli and apparatus Stimuli consisted of 160 commonly used Chinese four-word idioms and 44 nonwords. Most of these idioms were chosen from the online Chinese dictionary published by Ministry of Education, Taiwan (http://dict. revised.moe.edu.tw/cbdic/search.htm). The congruent condition consisted of the original sequences of the words in all the idioms, and the incongruent condition consisted of four-word sequences in which the ending of each idiom was substituted by a matched unrelated word. The matched words had the same stroke numbers and similar word frequencies (original vs. matched: 1,595 vs. 1,574), $t(159) = .153$, $p = .879$, as the originals. Also, to avoid any confounding from phonology, the matched words all had dissimilar consonants and vowels from the original ending words (see the Appendix in the supplementary materials). All nonwords were composed of different basic components of Chinese words that could not form real words.

Two viewing conditions (visible, invisible) were included to compare the role of consciousness in semantic temporal integration. In the invisible (i.e., CFS) condition, participants viewed dichoptically with the word ($1.03° \times 1.03°$) in their nondominant eye, and dynamic colorful Mondrians ($9.15° \times 9.15°$) changing every 100 ms (10 Hz) in their dominant eye. The Mondrians were made of small rectangles with various sizes, luminance, and colors (RGB values ranging from 0 to 255). In the visible condition, participants viewed binocularly with the words superimposed on the Mondrians in both eyes.

The contrast of the word, specifically for the invisible condition, was set according to a pretest before the actual experiment for each participant. The pretest began with 40% contrast and participants had to report whether they saw any word during CFS. The contrast increased 2.5% if the word was invisible and decreased 2.5% if any word was visible unless it reached the lowest level of 15%. The pretests would end after reaching six turning points or after 35 trials elapsed. The initial contrast of the invisible condition was the average contrast of the last 10 trials in the pretest and would decrease by 2.5% each time the word was reported to be visible, with the lowest contrast also set at 15%. In the visible condition, the contrast of the words remained at the contrast of 50% throughout.

Visual stimuli were presented to the participants using PsychoPy (Peirce, 2007), on an ASUS 22-inch VE228H LED monitor with a resolution of 1,280 × 1,024 pixels at 75 Hz refresh rate. Participants were seated at a distance of 100 cm in front of the monitor in a dark room, with his or her head stabilized on a chin-and-head rest. A set of four-mirror

stereoscopes was adjusted for each participant to ensure stable binocular fusion.

EEG recording and analysis EEG data were recorded with a 32-channel Quick-Cap using the Neuroscan System. The locations of scalp electrodes were placed according to the 10–20 system. Another two pairs of electrodes were placed above and below the left eye and at the outer canthus of both eyes to monitor vertical and horizontal eye movements. Electrode impedances were kept below 5 kΩ throughout the experiment. With SynAmps and a 0.05–200 Hz band-pass filter, all signals were amplified and continuously sampled at a 1000 Hz sampling rate. All EEG recordings were online referenced to a vertex reference electrode and off-line rereferenced to the average of the left and right mastoids.

EEG data were preprocessed off-line using the Neuroscan 4.5 software. EEG signals for trials with eye movements, blinks, body movements, muscle noise, and incorrect responses were excluded before epoching. EEG responses were time-locked to the onset of the targets—either a congruent or incongruent ending of the idioms or a nonword—and were epoched from −150 ms to 1,000 ms surrounding the onsets. EEG signals were digitally filtered by a low-pass filter of 30 Hz with the magnitude of 24 dB and baseline corrected to the interval of 150 ms before the stimulus onset. Epochs with signal amplitudes exceeding ±75 μV were removed as artifacts. The preprocessed ERPs were averaged for each participant on each stimulus type. Hence, we focused on the N400 time window (300–500 ms) for further grand average analysis.

Design In Experiment 1a, we adopted a 2 (viewing type: visible, invisible) × 2 (congruency: congruent, incongruent) × 2 (target: word, nonword) within-subjects design. Viewing types were counterbalanced between blocks, while the others were mixed within blocks. Each condition contained 40 trials, resulting in 320 trials in total. All idioms were presented once in each block in random order.

In Experiment 1b, we only tested the invisible condition with a doubled number of trials and with only behavioral measures, therefore, the design was changed to 2 (congruency: congruent, incongruent) × 2 (target: word, nonword) within-subjects design. All conditions were randomly intermixed. Each condition contained 80 trials, resulting in 320 trials in total, with each idiom presented twice in random order within each block.

Procedure During each trial in the invisible block (see Fig. 1a), the first three words of an idiom were sequentially presented for 250 ms each, followed by a fixation for a random duration ranging from 400 to 600 ms. An unmasked target was then presented after the fixation until response or 2,000 ms after the onset. Participants were instructed with the following:

> This is a lexical decision task. In each trial, you will see color masks first, then a target that is either a word or nonword. Please judge whether the target is a word or nonword as quickly and accurately as possible by pressing either the left- or right-arrow key, respectively. After the lexical decision judgment, please also judge whether or not you see any word before the target. There is no time pressure in this second response. Please try not to blink your eyes before the second response.

The keys for words and nonwords were counterbalanced between participants, and the second response served as a subjective visibility check. The next trial began after an intertrial interval ranging from 500 to 800 ms.

The visible block (see Fig. 1b) and the invisible block were identical except for the viewing type regarding how the first three words were presented. Participants viewed dichoptically in the invisible block and binocularly in the visible block, as described in the Stimuli and apparatus section. The order of visible and invisible blocks was counterbalanced between participants. Participants were allowed to take a brief break every 40 trials, and it took approximately 40 minutes to complete the experiment.

Results

Experiment 1a In the invisible block, the mean average contrast in the pretest (which was used as the initial contrast in the experiment) was 41.37% (range: 15%–80%), and the mean contrast in the experiment was 37.84% (range: 15%–80%). In the visibility check, the difference of mean visibility between the visible and invisible blocks was statistically significant (99.7% vs. 3.2%), $t(29) = 100.67$, $p < .001$. From this visibility check, we infer that the manipulation was effective in the sense that words in the invisible block were well-suppressed from consciousness. Those subjectively-reported invisible and visible trials were excluded from the visible and invisible block, respectively, from further analysis.

ERPs The ERP waveforms of the congruent and incongruent ending trials from electrode FP1, FP2, F7, F3, FZ, F4, F8, T7, C3, CZ, C4, T8, P7, P3, PZ, P4, P8, O1, and O2, were averaged across 30 participants in the invisible and visible block, respectively. In general, the ERPs in the incongruent ending trials were more negative than the congruent ending trials in the visible block within the 300–500-ms time window, a classical N400 effect of semantic violation (Kutas & Hillyard, 1980). However, this pattern was absent in the invisible block. To focus on the N400 component, the ERP amplitudes were averaged between a 300 and 500 time window across anterior
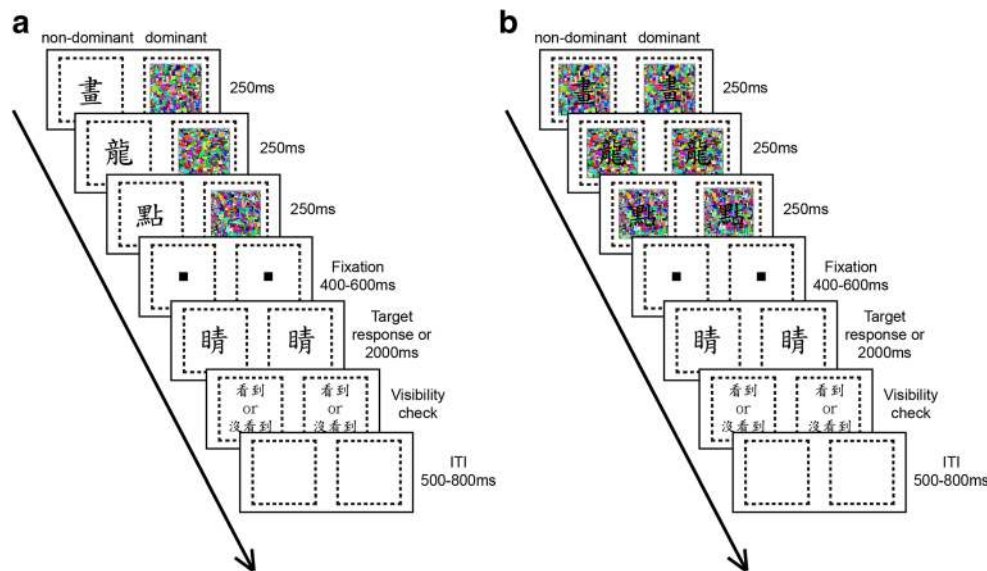
Fig. 1 The trial procedures in Experiment 1 (not to scale). **a** The trial procedure of the invisible block. **b** The trial procedure of the visible block. The first three words of an idiom were dichoptically (invisible block) or binocularly (visible block) presented word-by-words for 250 ms, followed by a fixation sign (400–600 ms), a target, and a visibility check. The target could be a correct ending of the idiom (the congruent condition), a matched but unrelated word (the incongruent condition), or a nonword. The example illustrated here was a congruent one with the correct ending. The idiom 畫龍點睛 literally means "draw a dragon and dot in the eye," which implies bringing a painting to life by adding the crucial touch

(F7, F3, Fz, F4, F8), central (T7, C3, Cz, C4, T8), and posterior (P7, P3, Pz, P4, P8) regions, and were submitted to a 2 (viewing type: visible, invisible) × 3 (region: anterior, central, posterior) × 2 (congruency: congruent, incongruent) repeated-measures ANOVA. Greenhouse–Geisser correction was adopted when Mauchly's test of sphericity was violated. The average waveforms of the anterior, central, and posterior regions were plotted in Fig. 2a, and the result showed significant main effects of viewing type, $F(1, 29) = 6.10$, $p = .02$, $\eta_p^2 = .174$;, region, $F(1.13, 32.71) = 10.70$, $p = .002$, $\eta_p^2 = .269$; and congruency, $F(1, 29) = 11.04$, $p = .002$, $\eta_p^2 = .276$. The two-way interactions of Viewing Type × Region, $F(1.14, 33.08) = 18.90$, $p < .001$, $\eta_p^2 = .395$; Viewing Type × Congruency, $F(1, 29) = 9.80$, $p = .004$, $\eta_p^2 = .253$; and Region × Congruency, $F(2, 58) = 13.04$, $p < .001$, $\eta_p^2 = .310$, were all significant. The three-way interaction of Visibility × Region × Congruency interaction was not significant ($p = .656$). LSD post hoc analysis demonstrated that N400 voltage of the incongruent endings were significantly more negative than the congruent endings only in the visible block: anterior, $t(29) = 2.57$, $p = .016$; central, $t(29) = 4.28$, $p < .001$; posterior: $t(29) = 5.29$, $p < .001$, but not in the invisible block: anterior, $t(29) = −1.004$, $p = .323$; central, $t(29) = −.266$, $p = .795$; posterior, $t(29) = .886$, $p = .383$. Furthermore, in order to verify the result that no N400 difference in the invisible block stands for the null hypothesis, we calculated the Bayes factors (Dienes, 2014) for the N400 data. The Bayes factors (BF_{10}) of the anterior, central, and posterior regions were 0.04, 0.00, and 0.00, respectively. All Bayes factors were much lower than one third, which strongly indicated that there was no N400 congruency effect in the invisible block (Lee & Wagenmakers, 2014). The difference voltages between the incongruent and congruent condition were shown in Fig. 2b, which emphasized a clear central-posterior N400 distribution in the visible block but not in the invisible block. The mean N400 amplitudes of the anterior, central, and posterior regions in the visible and invisible block are illustrated in Fig. 2c.

Behavioral data Since nonwords were adopted for the orthogonal task design (i.e., the LDT), only the trials with words (including correct and incorrect ending words) were used for accuracy and RT analysis in the current as well as the following experiments.

There was no difference in the overall accuracy of LDT between visible and invisible blocks (99.0% vs. 98.7%), $t(29) = 1.44$, $p = .161$, indicating that participants did not adopt any specific strategy between two different blocks. As shown in Fig. 3a, the accuracy data were submitted to a 2 (viewing type: visible, invisible) × 2 (congruency: congruent, incongruent) repeated-measures ANOVA. Results showed no main effects of viewing type, $F(1,29) = 3.95$, $p = .056$, $\eta_p^2 = .12$, and congruency, $F(1, 29) = 1.43$, $p = .242$, $\eta_p^2 = .047$. Their interaction was not significant, $F(1, 29) = 1.51$, $p = .229$, $\eta_p^2 = .05$. High accuracy (>98%) in all conditions might indicate a ceiling effect, leading to no congruency effects in accuracy between congruent and incongruent endings.

RTs of incorrect trials were excluded from further analysis. Correct RTs were submitted to a 2 (viewing type: visible,
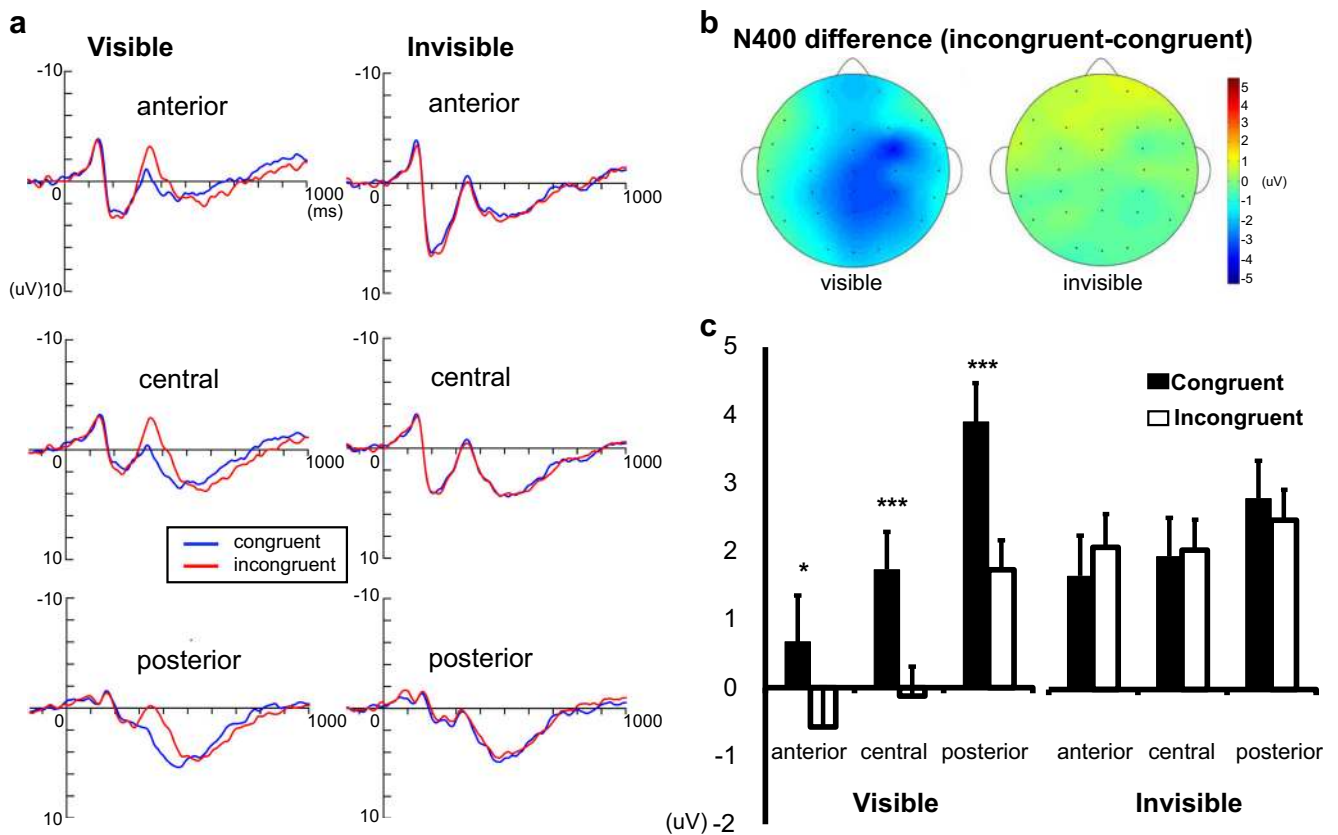
Fig. 2 N400 effects in Experiment 1a. **a** Mean ERP waveforms across the anterior (F7, F3, Fz, F4, F8), central (T7, C3, Cz, C4, T8), and posterior (P7, P3, Pz, P4, P8) regions. **b** Difference voltage map between 300 and 500 ms, which was calculated by the voltage of the incongruent condition minus that of the congruent condition. **c** Mean N400 amplitudes of the anterior (F7, F3, Fz, F4, F8), central (T7, C3, Cz, C4, T8), and posterior (P7, P3, Pz, P4, P8) among the invisible and visible blocks. N400s of the incongruent endings were significantly more negative than those of the congruent endings in the visible block, but no N400 difference was found in the invisible block. *Error bars* indicate the standard error of N400 amplitudes for each condition. $*p < .05$. $***p < .001$
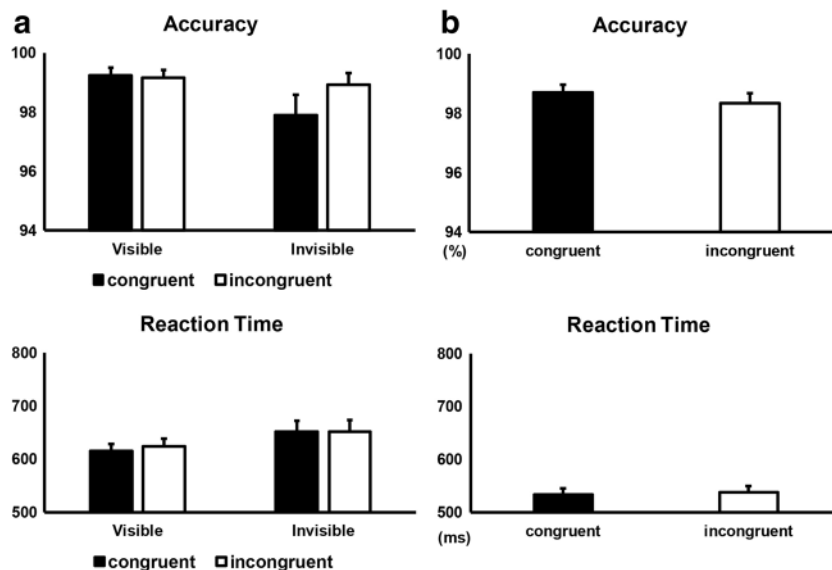

Fig. 3 Results of Experiment 1. **a** Mean accuracy and RT in Experiment 1a. Both accuracy and RT results showed no main effects of viewing type and congruency. **b** Mean accuracy and RT in Experiment 1b in which only the invisible type was included and the trial numbers were doubled. Both accuracy and RT results showed no difference between congruent and incongruent endings. *Error bars* indicate one standard error of the mean

invisible) × 2 (congruency: congruent, incongruent) repeated-measures ANOVA, which showed no main effects of viewing type, $F(1, 29) = 3.09$, $p = .089$, $\eta_p^2 = .096$, and congruency, $F(1, 29) = 2.68$, $p = .112$, $\eta_p^2 = .085$. The interaction was not significant, $F(1, 29) = 1.31$, $p = .261$, $\eta_p^2 = .043$. The planned comparison illustrated that the mean RTs for the incongruent endings were similar to the congruent endings in the visible block (623.6 ms vs. 614.1 ms), $t(29) = 1.99$, $p = .056$, and in the invisible block (652.0 ms vs. 651.8 ms), $t(29) = .22$, $p = .828$. There was no speed–accuracy trade-off.

Experiment 1b The initial contrast of the experiment was 59.67% (range: 15%–88%), and the mean contrast in the experiment was 50.93% (range: 15%–85%). The mean visibility in Experiment 1Bb was 4.6% ($SD = 7.42\%$), and all visible trials were removed from further accuracy and RT analysis.

The overall accuracy of the LDT was 98.2%, and there was no accuracy difference between the congruent and incongruent endings (98.7% vs. 98.3%), $t(29) = 1.38$, $p = .177$. RT analysis of the correct trials also showed no difference between the congruent and incongruent endings (533.5 ms vs. 538.3 ms), $t(29) = -.64$, $p = .526$, which suggested that there was no speed–accuracy trade-off. All behavioral results in Experiment 1 are illustrated in Fig. 3.

Discussion

Our analyses in the current experiment yielded the finding that congruent endings led to smaller N400 responses than incongruent endings in the visible condition. Such finding indicated that the first three words of the idioms could be temporally integrated into idiom contexts, which replicated the previous study using Chinese idioms (Liu et al., 2010). However, no such temporal integration occurred when the preceding words were suppressed by CFS. This result was replicated even with enhanced statistical power in Experiment 1b, which exhibited no accuracy and RT differences between congruent and incongruent endings. The absence of N400 and RT congruency effect in the invisible condition suggests that semantic temporal integration did not occur when words were not consciously perceived. However, one might argue that the null results in Experiments 1a and 1b were due to lack of top-down attention, since top-down attention is important for subliminal processing to occur (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). Thus, in the next experiment, we investigated whether subliminal words could be integrated into idioms with enhanced top-down attention.

Experiment 2

In Experiments 2a and 2b, we further probed if the subliminal word sequences could be integrated into idioms with

enhanced top-down attention through both behavioral and EEG measures. Top-down attention is thought to play an important role in subliminal processing (Dehaene et al., 2006), and a study has demonstrated that top-down attention might serve as a prerequisite for subliminal semantic priming to occur (Van den Bussche, Hughes, Humbeeck, & Reynvoet, 2010). To enhance top-down attention during the presumed integration process, we added a white fixation frame at the corresponding location of the masked words on the Mondrians and asked participants to focus on the center of the frame. Participants were required to detect and report if any words were presented in the frame while performing the LDT. If the integration does not require consciousness but requires top-down attention, the congruency effect—smaller N400s—would be found in the congruent condition when compared with the incongruent condition. Otherwise, if no congruency effect is found, it would suggest that semantic temporal integration does not occur unconsciously regardless of top-down attention manipulations.

In Experiment 2a, besides attention enhancement, in order to avoid possible behavioral response artifacts at the aimed epoch (i.e., congruent and incongruent endings, the focus of the N400 analysis), we changed the task into a go/no-go LDT. Participants were asked to press a key when the nonword targets appeared but told to withhold response if the word targets appeared instead. This change allowed us to ensure artifact-free ERPs, which also enhanced the sensitivity of detecting the semantic congruency effect. However, since no RT responses were collected under this setting, another group of participants was tested with the same setup (top-down attention enhanced) under the original LDT task without EEG recordings in Experiment 2b.

Method

Participants Twenty-four (18–27 years) participants were recruited in Experiment 2a and 30 (18–29 years) in Experiment 2b.

Stimuli and apparatus All stimuli and apparatus were the same as in Experiment 1, with the exception that a white fixation frame (1.06° × 1.06°) was superimposed on the corresponding location of the masked words on the Mondrians.

EEG recording and analysis EEG recording and analysis were the same as in the invisible condition of Experiment 1.

Design A 2 (congruency: congruent, incongruent) × 2 (target: word, nonword) within-subjects design was conducted, with all four types of trials randomly intermixed. Each condition contained 80 trials, resulting in 320 trials in total, with each idiom presented twice in random order.

Procedure A trial began with the first three words of an idiom sequentially presented for 250 ms in the nondominant eye, while dynamic colorful Mondrians were flashed in the dominant eye with a central fixation frame. A fixation was then binocularly presented for a random duration ranging from 400 to 600 ms, followed by a binocularly presented unmasked target. The target was either a congruent ending, an incongruent ending, or a nonword. In Experiment 2 s, Participants were instructed with the following:

> This is a lexical decision task. In each trial, you will see color masks with a white frame. Please focus on the middle of the frame and detect whether any words were presented in the frame or not. After that, you will see a target. Please press the left-arrow key as quickly and accurately as possible if you see a nonword target, and please hold on when it is a word target. After the lexical decision judgment, please also judge whether you saw any word before the target in the frame. There is no time pressure in this second response. Please try not to blink before the second response.

In Experiment 2b, participants were instructed similarly to Experiment 2a but were asked to "Please judge whether the target is a word or non-word as quickly and accurately as possible by pressing either the left- or right-arrow key" as in Experiment 1. The next trial began after an intertrial interval ranging from 500 to 800 ms (see Fig. 4a).

Results

Experiment 2a The mean initial contrast was 45.08% (range: 15%–80%), and the mean contrast in the experiment was 30.89% (range: 15%–80%). The overall mean accuracy of the go/no-go LDT was 98.83%, and the accuracy in the congruent endings (99.29%) and incongruent endings (99.46%) did not differ significantly, $t(23) = -.624$, $p = .539$, suggesting that all participants performed well in the task. The visibility check yielded a mean visibility of 7.6% ($SD = 12.4\%$). All visible trials were excluded from further ERP analysis. As shown in Fig. 4b–d, a 3 (region: anterior, central, posterior) × 2 (congruency: congruent, incongruent) repeated-measures ANOVA found no main effects of region, $F(1.13, 25.98) = .01$, $p = .934$, $\eta_p^2 = .001$, and congruency, $F(1, 23) = .001$, $p = .973$, $\eta_p^2 = .000$. The Region × Congruency interaction was not significant, $F(2,46) = 1.41$, $p = .254$, $\eta_p^2 = .058$. Planned comparisons also indicated the absence of congruency effect as revealed by the N400 component: anterior, $t(23) = .44$, $p = .662$; central, $t(23) = -.08$, $p = .941$; posterior, $t(23) = -.392$, $p = .698$, which yielded Bayes factors (BF_{10}) of 0.11, 0.01, and 0.01, respectively. Our results indicated that there was no N400 congruency effect when the words were invisible.

Experiment 2b The mean initial contrast was 55.23% (range: 15%–80%), and the mean contrast in the experiment was 42.51% (range: 15%–80%). The mean visibility of the subjective visibility check was 6.4% ($SD = 8.3\%$), and all visible trials were excluded from further accuracy and RT analysis. The overall accuracy of the LDT was 98.4%, and there was no accuracy difference between the congruent and incongruent endings (98.1% vs. 98.1%), $t(29) = .20$, $p = .845$. The analysis of the correct RT trials also indicated no significant difference between the congruent and incongruent endings (605.8 ms vs. 608.6 ms), $t(29) = -.74$, $p = .468$, which means that there was no speed–accuracy trade-off. All results in Experiment 2 are presented in Fig. 4.

Discussion

In this experiment, we enhanced top-down attention to increase the possibility of temporal integration of semantic information. However, we found no N400 amplitude and RT differences between the congruent and incongruent endings, which was consistent with Experiment 1. This suggests that semantic temporal integration does not occur unconsciously, even with top-down attention enhancement.

One possible reason that we observed the null results of invisible condition in Experiments 1 and 2 could be due to the long ISIs between the prime and the target. Since unconscious processing may be short lived, the ISIs between prime and target in previous experiments might be too long (400–600 ms) to allow for temporal integration. For example, Kiefer and Spitzer (2000) showed that N400 is modulated by masked words only when stimulus onset asynchrony (SOA) between prime and target was short (67 ms) but not long (200 ms). We therefore excluded the ISI between prime and target in the next experiment to see whether temporal integration can be observed.

Experiment 3

Similar to Experiment 2a, a go/no-go LDT task was used in conjunction with EEG recordings in this experiment. In addition to using no ISI in this experiment to facilitate the possibility of temporal integration, we also aimed to exclude one possible confounding factor in the previous experiments: contrast-dependent unconscious processing. ERP results in Experiment 1a showed temporal integration in the visible condition but not in the invisible condition. The conclusion that consciousness is required for temporal integration, however, seems premature without having the results confirmed with equivalent contrast, since the contrast was higher in the visible condition than the invisible condition in Experiment 1a. Previous studies have shown that contrast of prime affects the priming effect (Tapia & Breitmeyer, 2011); it is thus
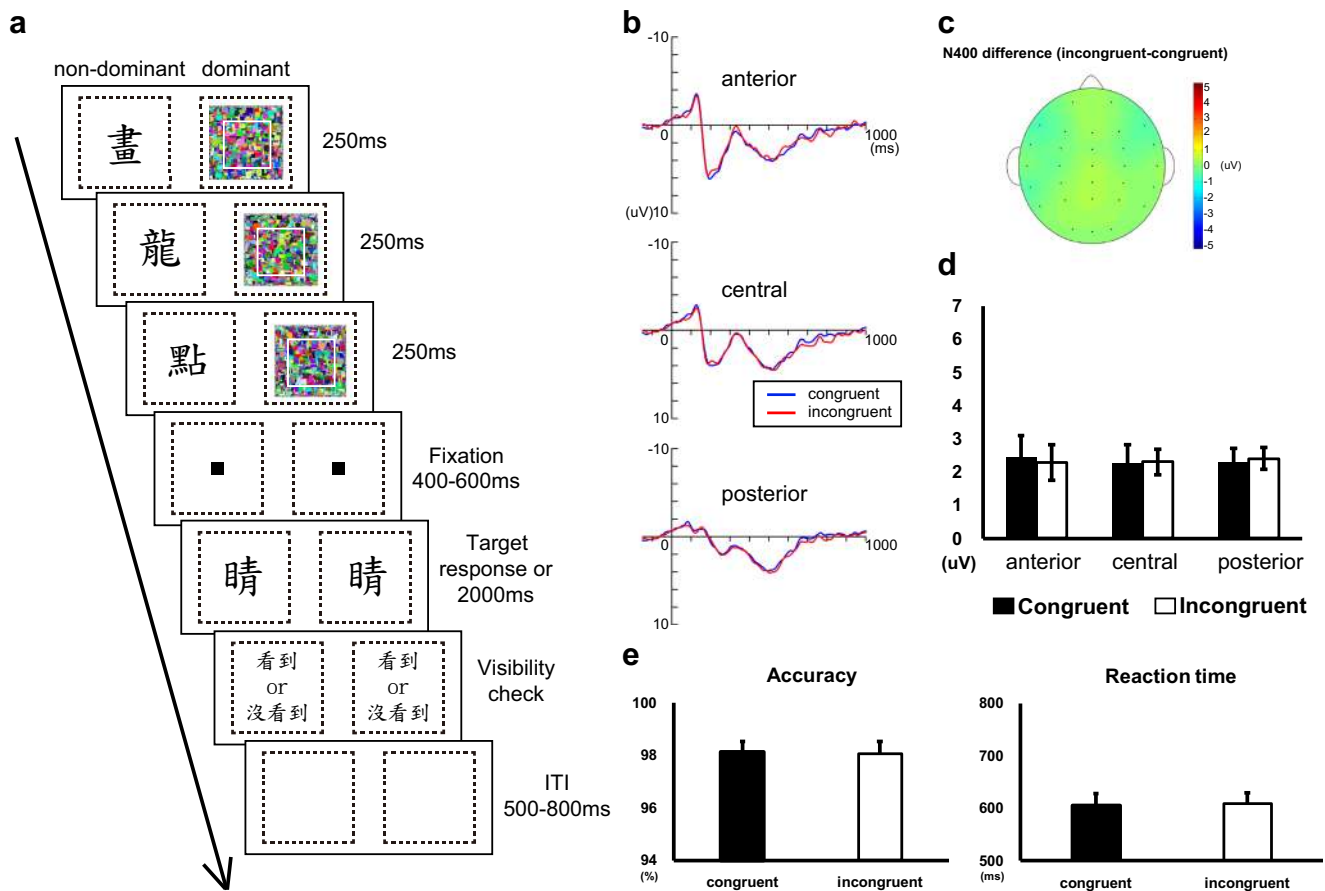
Fig. 4 Procedure and results in Experiment 2. **a** Trial procedure in Experiment 2 (not to scale). **b** Experiment 2a: Mean ERP waveforms across the anterior (F7, F3, Fz, F4, F8), central (T7, C3, Cz, C4, T8), and posterior (P7, P3, Pz, P4, P8) regions. **c** Experiment 2a: The difference voltage map between 300 and 500 ms, which was calculated by subtracting the voltage of the congruent condition from the voltage of the incongruent condition. However, the difference between the two conditions was not significant. **d** Experiment 2a: Mean N400

amplitudes of the anterior (F7, F3, Fz, F4, F8), central (T7, C3, Cz, C4, T8), and posterior (P7, P3, Pz, P4, P8) regions. Again, no N400 difference was found between the congruent and incongruent endings. *Error bars* indicate the standard error of the N400 amplitudes for each condition. **e** Mean accuracy and RT results in Experiment 2b, both showed no difference between the congruent and the incongruent endings. *Error bars* indicate the standard error of the mean

important to dissociate the covariant contribution of the contrast of stimuli and the conscious state to the priming effect. We therefore used the same contrast in the superimposed and CFS conditions in the current experiment.

Method

Participants Another group of 18 participants (19–31 years old) took part in this experiment.

Stimuli, apparatus, and procedure All methods were the same as in Experiment 1a, except for the following. (1) The ISI between the third word and the fourth word was zero ms (i.e., no ISI between prime and target). (2) The contrast of words was the same in both superimposed and CFS conditions. (3) Unlike in Experiment 1, in which the adaptive contrast method was used depending on the visibility, fixed contrast was used in the whole experiment here. (4) Similar to previous experiments, a pretest

was used to find a suitable contrast for each participant before the experiment, but with a modified method: the four-up-one-down staircase method (i.e., the contrast increased 2.5% if the word was invisible four times and decrease 2.5% if any word was visible once); this would make the probability of visibility at around 15.9% (Levitt, 1971). (5) The size of Mondrians became 2° × 2° on the corresponding location of masked word, which could directly constrain top-down attention on the word position. (6) Participants were instructed to do the go/no-go LDT task as in Experiment 2a to ensure artifact-free ERPs.

The stimuli and procedure were conducted using MATLAB Version 2012a (the Math Works, Natick, USA) with Psychophysics Toolbox 3.0 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

Design We adopted a 2 (viewing type: visible, invisible) × 2 (congruency: congruent, incongruent) × 2 (target: word, non-word) within-subjects design. Viewing types were

counterbalanced between blocks, while the other factors were mixed within blocks. Each condition contained 40 trials, resulting in 320 trials in total, with each idiom presented once in each block in random order.

Results

The mean initial contrast was 53.94% (range: 20%–86%), and we used this contrast throughout the experiment in both the visible and invisible conditions. The mean visibility of the subjective visibility check in the CFS condition was 20.9% ($SD = 25.9\%$), and all visible trials were excluded from further ERP analysis.

Behavioral data We conducted a 2 (viewing type: visible, invisible) × 2 (congruency: congruent, incongruent) repeated-measures ANOVA on accuracy. There was neither main effect of viewing type, $F(1, 17) = .079$, $p = .782$, $\eta_p^2 = .005$, nor main effect of congruency, $F(1, 17) = .000$, $p = 1$, $\eta_p^2 = .000$, nor interaction between viewing type and congruency, $F(1, 17) = .654$, $p = .430$, $\eta_p^2 = .037$.

ERPs data As shown in Fig. 5, We focused on the N400 component (300–500 ms) with averaging regions of anterior (F7, F3, Fz, F4, F8), central (T7, C3, Cz, C4, T8), and posterior (P7, P3, Pz, P4, P8), and submitted to a 2 (viewing type: visible, invisible) × 3 (region: anterior, central, posterior) × 2 (congruency: congruent, incongruent) repeated-measures ANOVA. The main effect of congruency was significant, $F(1, 17) = 6.616$, $p = .02$, $\eta_p^2 = .280$. The Region × Congruency interaction, $F(2, 34) = 5.156$, $p = .011$, $\eta_p^2 = .233$, and Viewing Type × Congruency interaction, $F(1, 17) = 7.235$, $p = .016$, $\eta_p^2 = .299$, were significant. LSD post hoc analysis showed that N400 component of the incongruent endings were significantly more negative than the congruent endings only in the visible condition: anterior, $t(17) = 2.703$, $p = .015$; central, $t(17) = 4.045$, $p = .001$; posterior, $t(17) = 4.56$, $p < .001$, but not in the invisible condition: anterior, $t(17) = -.851$, $p = .407$; central, $t(17) = .490$, $p = .631$; posterior, $t(17) = .810$, $p = .429$, which yielded Bayes factors (BF_{10}) of 0.29, 0.07, and 0.09, respectively.

Discussion

In this experiment, we tested whether temporal integration could be found in the invisible condition when there was no ISIs between prime and target words. However, N400 effect was still not found in the invisible condition. This result implies that a long duration between prime and target was not the main reason for the absence of priming effect in Experiments 1 and 2.

More specifically, we used the same contrast between the visible and invisible conditions to rule out the possible confounding of contrast-dependent priming effect and the result

was similar to studies that showed stronger priming effects for the visible condition than that for the invisible condition (Tapia & Breitmeyer, 2011; Tapia, Breitmeyer, & Shooner, 2010), even after adopting the same contrast. However, unlike these studies that still showed a weak priming effect in the invisible condition, the N400 component was absent in the current experiment. One major reason could be that multiple words were needed to integrate across time, and consciousness is required for semantic integration of multiple words in the temporal domain.

Experiment 4

Previous studies of unconscious multisensory integration suggested that conscious training is essential for unconscious integration to occur (Faivre, Mudrik, Schwartz, & Koch, 2014; see also Noel, Wallace, & Blake, 2015). For example, Faivre et al. (2014) found that semantic-based visual-auditory integration could occur in complete unawareness of any sensory input, but only when the participants had been consciously trained. Therefore, in Experiment 4a, in order to enhance unconscious processing, we always conducted the visible condition before the invisible condition as a conscious training (Fig. 6)a–b.

Moreover, since previous studies showed that human can unconsciously integrate spatial information (Sklar et al., 2012), we tested whether the semantic information of multiple words can be unconsciously integrated in spatial domain (Experiment 4b). To analogize the temporal integration of a multiple word of LDT task in a spatial domain, we simultaneously presented the first three words at different locations, and the fourth stimulus was sequentially presented at the neighboring location of the third word (see Fig. 7)a–b.

Method

Participants Twenty-five (18–36 years old) and 27 naïve volunteers (20–30 years old) took part in Experiment 4a (temporal integration) and Experiment 4b (spatial integration), respectively.

Stimuli and apparatus In Experiment 4a, we tested temporal integration of semantic information. All the stimuli were the same as in Experiment 3, except that Mondrians were excluded in the visible condition for better conscious training (see Fig. 6a–b). In Experiment 4b, we tested spatial integration of semantic information. Participants dichoptically viewed the three words in their nondominant eye, and dynamic colorful Mondrians (9.15° × 9.15°) were changed every 100 ms (10 Hz) in their dominant eye. The three words (6° × 2°) were arranged spatially from left to right. The center-to-center distance from frame to words were left 1.5°, left 0.5°, and right
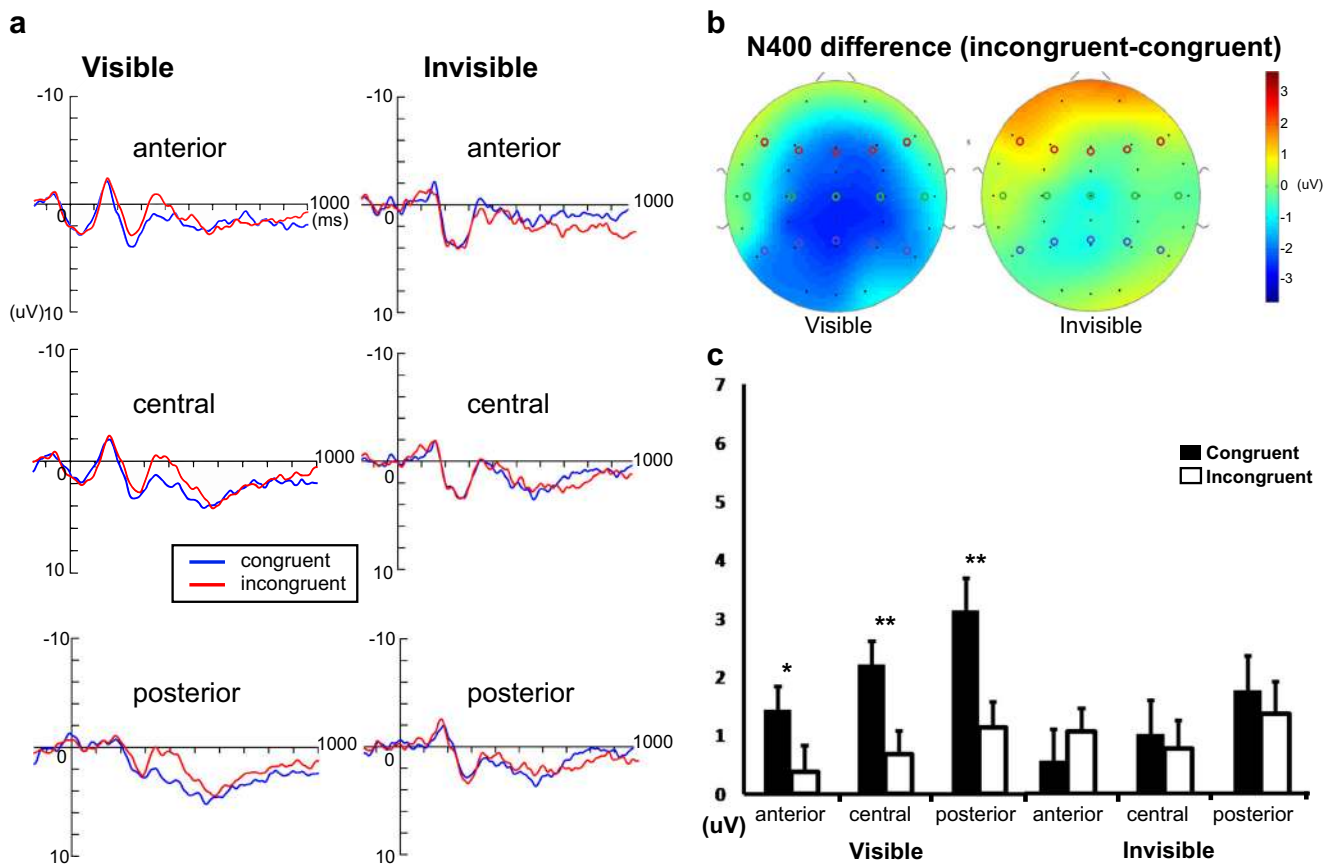
Fig. 5 The N400 effects in Experiment 3. **a** Mean ERP waveforms across the anterior, central, and posterior regions. **b** Brain topography of the N400 difference. **c** Mean N400 amplitudes of the anterior, central, and posterior among the invisible and visible blocks. N400s of the incongruent endings were significantly more negative than those of the congruent endings in the visible block, but no N400 difference was found in the invisible block. *Error bars* indicate the standard error of N400 amplitudes for each condition. $*p < .05$. $**p < .01$

$0.5°$, respectively. The fourth target ($2° \times 2°$) was either congruent word, incongruent word, or nonword, as in previous experiments, and presented to the right $1.5°$ from the center of the frame (see Fig. 7a–b). Fixed contrast of the word (30%) was used throughout the experiment.

Design In both Experiments 4a and 4b, we adopted a 2 (viewing type: visible, invisible) × 2 (congruency: congruent, incongruent) × 2 (target: word, nonword) within-subjects design. The visible and invisible conditions were conducted sequentially, and the other factors were mixed within blocks. Each condition contained 80 trials, resulting in 640 trials, with each idiom presented four times in random order.

Procedure The LDT task and the instruction were the same as in Experiment 2b. The rest of the procedure in Experiment 4a (see Fig. 6) was the same as in Experiment 3. In Experiment 4b, the first three words were simultaneously presented for 750 ms, and then the target stimulus was presented for 2,000 ms. After that, participants were required to report their visibility of the prime words (see Fig. 7).

Results

Since the visible condition was always conducted before the invisible condition as conscious training to facilitate unconscious processing in the invisible condition (Faivre et al., 2014), we analyzed visible and invisible condition separately since presentation order is covaried with visibility. Participants were removed for having visibilities higher than 50% in the invisible condition and/or accuracies lower than 80% in the LDT task.

Experiment 4a The mean visibility of the subjective visibility check in the CFS condition was 5.0% ($SD = 9.5\%$), and all visible trials were excluded from further analysis in the invisible condition. One participant was excluded for further analysis due to high visibility (mean = 97.1%) in the invisible condition. In the visible condition, the mean accuracy revealed significant difference between congruent and incongruent endings (98.0% vs. 96.0%), $t(23) = 3.24$, $p = .004$. However, there was no significant difference in correct RT data (554.4 ms vs. 566.2 ms), $t(23) = -.954$, $p = .350$. In the
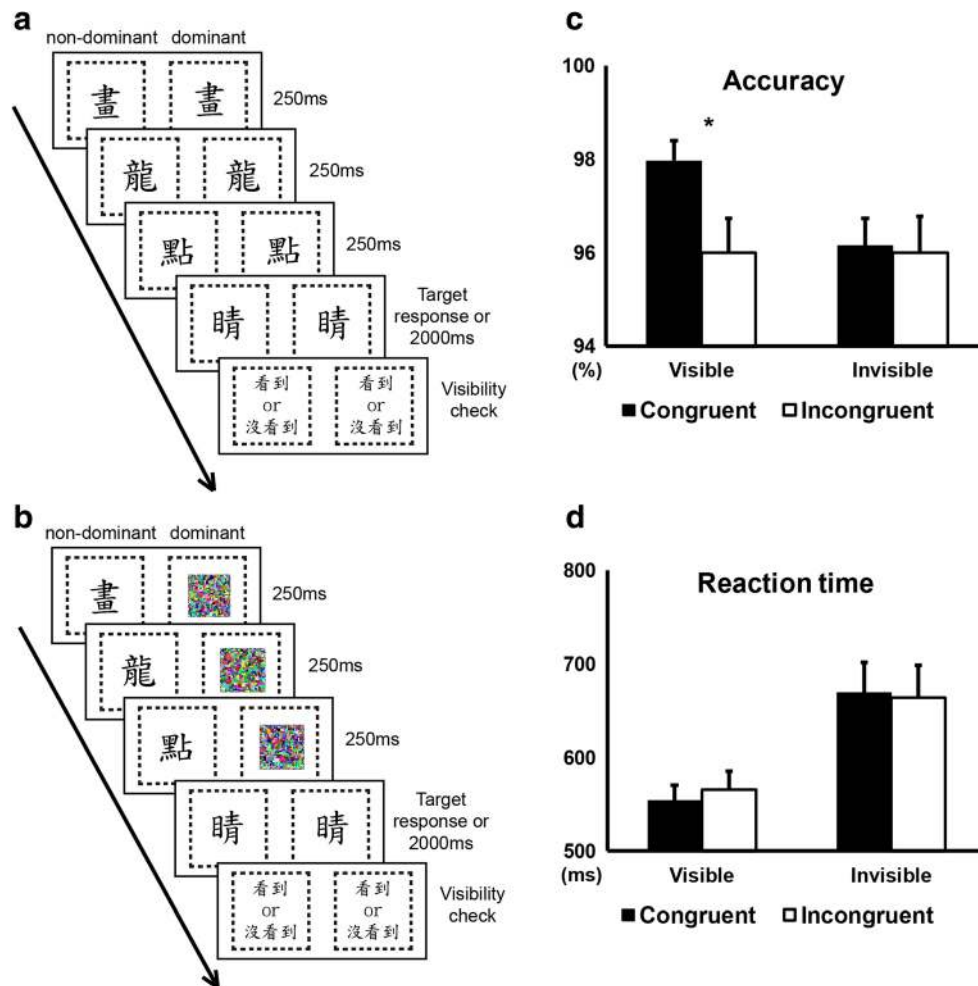
Fig. 6 The procedure and results in Experiment 4a. **a** Trial procedure of the visible block. **b** Trial procedure of the invisible block. **c** Mean accuracy revealed significant difference between congruent and incongruent endings in the visible condition. **d** Mean RT showed no congruency in both visible and invisible condition. *Error bars* indicate the standard error of the mean

invisible condition, we did not observe any difference between congruent and incongruent endings in accuracy (96.2% vs. 96.0%), $t(23) = .224$, $p = .825$, and correct RT (668.9 ms vs. 663.6 ms), $t(23) = .394$, $p = .697$, data (see Fig. 6c–d).

Experiment 4b The mean visibility of the subjective visibility check in the CFS condition was 9.7% ($SD = 16.2\%$), and all visible trials were excluded from further analysis in the invisible condition. Three participants were excluded for further analysis: two due to high visibility (mean = 86.97%) in the invisible condition, and one due to low accuracy (34.58%). In the visible condition, accuracy between the congruent and incongruent endings reached significant difference (97.4% vs. 95.6%), $t(23) = 3.436$, $p = .002$. The correct RT results also showed that RTs for congruent endings were shorter than for incongruent endings (509.6 ms vs. 563.0 ms), $t(23) = -4.465$, $p < .001$. In the invisible condition, the results showed that neither accuracy (96.67% vs. 97.24%), $t(23) = −1.161$, $p = .257$,

nor correct RT (689.4 ms vs. 693.0 ms), $t(23) = −.321$, $p = .751$, reached significant difference between congruent and incongruent endings in the invisible condition (see Fig. 7c–d).

Discussion

In this experiment, we conducted the visible condition before the invisible condition as conscious training to enhance unconscious processing. However, we still found that congruency effect was only revealed in the visible condition, and not in the invisible condition. The behavioral results in Experiment 4a were consistent with the N400 effect in Experiment 3. These also imply that consciousness is required for semantic integration of multiple words in the temporal domain, and that conscious training is not sufficient for unconscious temporal integration.

We also conducted the experiments with semantic integration in the spatial domain. Both accuracy and RT showed that spatial integration of semantic information could be processed
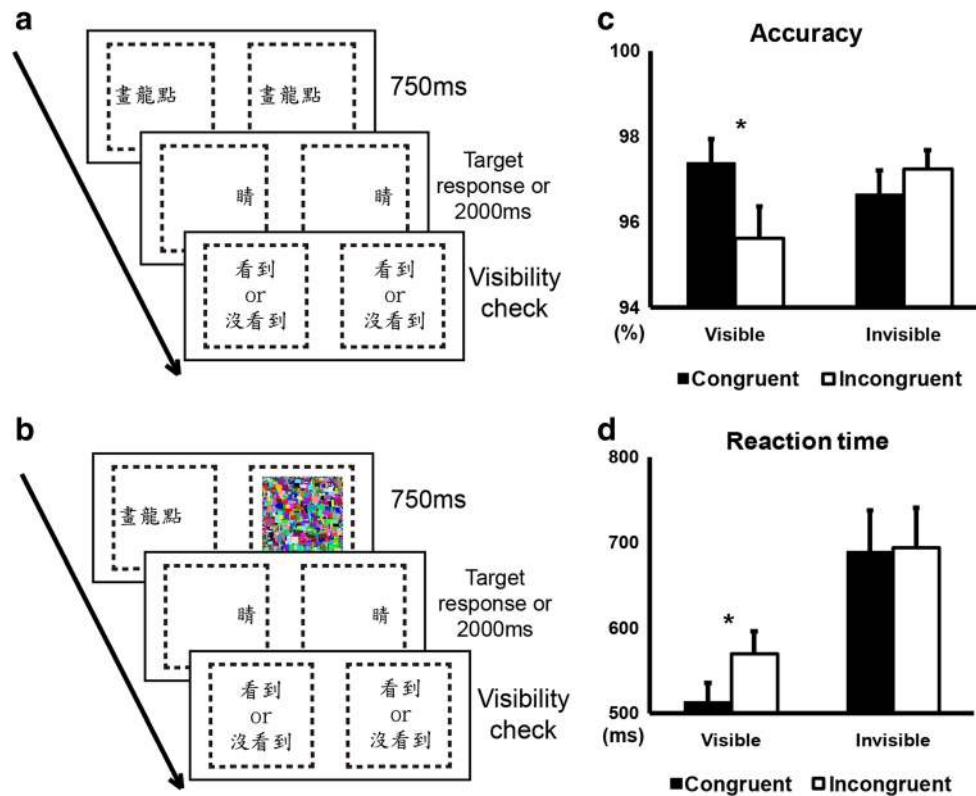
Fig. 7 Procedure and results in Experiment 4b. **a** Trial procedure of the visible block. **b** Trial procedure of the invisible block. **c** Mean accuracy revealed significant difference between congruent and incongruent endings in the visible condition. **d** Mean RT also showed congruency effect in the visible condition. There were no other significant differences. *Error bars* indicate the standard error of the mean

in the visible condition. However, different from what we expected, the congruency effect was not found in the invisible condition. It is possible that while the first three words were presented simultaneously, temporally integrating their semantic information with the fourth target word was still necessary. In other words, it still required spatial-temporal integration to induce the semantic priming effect.

Experiment 5

In previous experiments, we tested whether semantic information could be unconsciously integrated by adopting the LDT task. The basic assumption of this task was that the first three words can unconsciously form a semantic context and influence the discrimination of subsequently presented conscious target words. However, the LDT task in Experiment 4b required spatial-temporal integration that may have been too difficult to process unconsciously. In the current experiment, we tested whether semantic information can be integrated from multiple words if both spatial and temporal domain coincided.

Instead of using LDT task as in previous experiments, in this experiment we adopted breaking CFS (called b-CFS hereafter; Stein, Hebart, & Sterzer, 2011) task. In this task, the dynamic Mondrians were also presented to one eye as in

previous experiments, but the contrast of the suppressed stimulus presented to the other eye increased over time. Consequently, at some point the dominant percept would switch from Mondrians to the originally suppressed stimulus, and the time required for the suppressed stimulus to break CFS so as to gain access to consciousness was measured (Jiang, Costello, & He, 2007). Note that unlike the priming effect under CFS (e.g., Experiments 1–4), which indirectly measured the effect of fully invisible stimuli on a subsequent target, the b-CFS time directly measured the time for subjective visibility of suppressed stimulus. However, the b-CFS time can only be an index of how efficiently a suppressed stimulus is accessed into visual awareness, and whether it can reflect purely unconscious processing of suppressed stimuli is still debated (Gayet, Van der Stigchel, & Paffen, 2014; Stein & Sterzer, 2014). By adopting a similar procedure as in Sklar et al. (2012), we simultaneously presented four words side by side and asked participants to detect any part of the word sequence under CFS. To better compare the semantic integration of idioms, we included random sequence and inverted idioms as baselines.

Method

Participants Twenty-seven healthy naïve volunteers (20–30 years old) took part in the experiment.

Stimuli and apparatus Each frame contained four words to one eye, which were interocularly suppressed by colorful masks (9.15° × 9.15°) changing every 100 ms (10 Hz) to the other eye. The four words (6° × 2°) were arranged side by side from left 1.5°, left 0.5° to right 0.5°, right 1.5°, and either above 1° or below 1° from the word center to frame center (see Fig. 8a). The words from the same 160 idioms used previously were repeated three times: once as original idioms, once completely randomized and mixed with words from other idioms, and once with an inverted presentation (upside down) of idioms. The contrast of words was ramped from 0 to 100% within 1 second to prevent a sudden onset triggering a percept of the word in the beginning of the trial. The contrast of words was then kept at 100% after 1 second until the end of the trial.

Design This experiment was a one factor with three levels within-subjects design, which included idioms, random, and inverted conditions. Each condition contained 160 trials, resulting in 480 trials, with each idiom presented three times in random order.

Procedure Participants were instructed as follows:

> Please press the "z" key to start each trial. After that, you will see color masks and we will present words above or below the center. Your task is to detect any part of those words and judge their location. Please press "o" if you see that the words are presented above the center, and "k" if below. Please respond as quickly and accurately as possible, and do not blink during the trial. You can take a brief break before you press the "z" key to start the next trial.

Each trial terminated when the participant responded; otherwise, the trial would stop after 6 seconds if no response was made.

Results

The mean accuracy of location judgment was 98.6% (SD = 1.56%), and only correct trials were included for further analysis. Participants were removed for having accuracies lower than 80% in the location judgment task. Based on this criteria, two participants were excluded from further analysis due to low accuracy (mean = 48.13%, SD = 3.24%). We conducted a one-way within-subjects ANOVA; the main effect of correct RT reached significance, $F(2, 48) = 9.062$, $p < .001$, $\eta_p^2 = .274$. Post hoc analysis showed that idioms were detected faster than inverted words (1.848 s vs. 1.890 s), $t(24) = -3.030$, $p = .006$; and random words (1.8391 s) were also detected faster than inverted words, $t(24) = -3.753$, $p = .001$. There was no significant difference between idioms and random words, $t(24) = .837$, $p = .411$ (see Fig. 8b).

Discussion

In this experiment, the inverted condition was detected slower than the idiom and random conditions, which is in part consistent with our previous studies that showed the word inversion effect (Yang & Yeh, 2011, 2014). The inversion effect suggests that familiarity of words (i.e., the upright orientation) can be processed under CFS, which is also consistent with the well-established inversion effect under CFS, although with different stimulus categories (Jiang et al., 2007; Stein, Sterzer, & Peelen, 2012); yet the result of no difference in RTs between the idiom and random conditions implies that semantic information of multiple words cannot be integrated even with spatial-temporal coincidence.

Our result here is different from that in Sklar et al. (2012). They manipulated three-word sentences in Hebrew to form either semantic violation (e.g., "I broke the water") or semantic nonviolation (e.g., "I heated the water") sentences using the b-CFS task, and found that semantic violation sentences led to
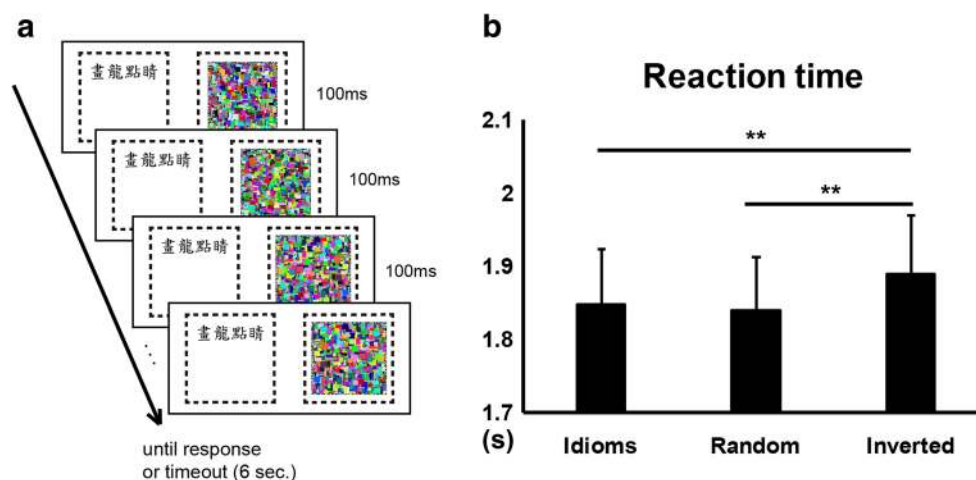

Fig. 8 Procedure and results of Experiment 6. **a** Trial procedure of the b-CFS task. **b** Mean RT showed that inverted words were detected slower than idioms words and random words. *Error bars* indicate the standard error of the mean

shorter b-CFS time than their counterparts. However, we did not observe such unconscious spatial integration, and several possible reasons could have induced this discrepancy: First, unlike in Sklar et al. (2012), where three words were used to form semantic meaning, four-word idioms were adopted in the current study. The additional word may have reached the capacity limit of semantic integration under CFS. Second, Sklar et al. (2012) used the typical subject-verb-object (SVO) sentence structure as stimuli, and the word that formed semantic violation always occurred in a fixed position of the sentence (e.g., verb in their Experiment 1 and subject in their Experiment 2). On the other hand, the Chinese idioms we used generally included subject, verb, object, and complement without fixed position, and this complex (and more ecological) structure may have increased the difficulty in semantic integration. Third, the semantic violation sentences comprised only one fourth of the total trials in their case; this manipulation could probably provide unequal expectation (if any) between semantic violation and semantic nonviolation sentences. Consequently, the shorter b-CFS time of semantic violation might be due to the "surprised" (in their term, p. 19617) feeling because of the less likely occurrence of the semantic violation sentences. Finally, unlike Hebrew, which is an abjad writing system, Chinese is a logographic writing system. The close orthographic-semantic relationship in Chinese could speed up the semantic processing compared to Hebrew. If semantic meaning of each word was unconsciously processed faster than semantic integration of all words, the task "to detect any part of the words" may have only needed the unconscious semantic processing of any particular word without semantic integration. Therefore, compared to the inverted condition, since both the idiom condition and the random condition preserved word orientation (both upright) and thereby the meaning of each word, a ceiling effect due to semantic processing of each word could induce similar reaction times between these conditions.

General discussion

This study examined whether semantic information could be temporally integrated without consciousness. In Experiment 1a, we compared the possibility of temporal integration of Chinese four-word idioms under visible and invisible blocks by presenting the first three words of Chinese four-word idioms sequentially, either binocularly with the Mondrians (the visible block) or dichoptically under CFS (the invisible block). Smaller N400 amplitude of the congruent endings than the incongruent ones in the visible condition was found, showing that temporal integration is possible when each word was visible and providing a baseline measure for the sensitivity of the paradigm used here. However, the congruency effect reflected by the ERP N400 effect and the behavioral measure

disappeared in the invisible block (Experiment 1a), and this is true even with doubling the trials by concentrating on the invisible block with the behavioral experiment to enhance statistical power (Experiment 1b). Moreover, we added a fixation frame to increase top-down attention in Experiment 2, and the congruency effect still could not be observed under CFS with both ERP and behavior measurements. In Experiment 3, the N400 effect was again found in the visible condition but not in the invisible condition when we removed prime-target ISI to overcome the potentially short-lived unconscious processing and used the same contrast in the visible and invisible conditions to exclude the possible confounding of contrast-dependent priming effect. Experiment 4 added a preceding conscious training to enhance unconsciously semantic integration, and the results still showed no unconscious semantic integration. Results of Experiment 5 indicated that the absence of unconscious semantic integration was not due to the insensitivity of our paradigm by showing that familiarity of word orientation could be unconsciously processed under CFS under the current setting. Overall, our results suggest that while consciousness is not necessary for semantic processing, it is required for semantic integration of multiple words (see Table 1 for a summary).

Our findings of an N400 effect in the visible condition but not in the invisible conditions cannot be simply explained by the reason that N400 is an index only for conscious integration of semantic violation since several studies revealed that unconscious words can also modulate the N400 component (Kiefer, 2002; Kiefer & Spitzer, 2000; van Gaal et al., 2014). For example, Kiefer (2002) used masked single words and found that semantically unrelated words revealed a larger N400 effect than related words. That is, the N400 effect can also be an index of semantic violation from the invisible prime. More specifically, Heyman and Moors (2012) argued that the null finding of N400 effect while using interocular suppression in combination with EEG must consider the possibility of delayed N400 as well as the nature of N400 effect (see also Kang et al., 2011). Nevertheless, such delayed N400 might still be insufficient to fully explain the absence of the N400 congruency effect in the invisible condition of the current study. First, the delayed N400 was mostly found with unidentifiable target stimuli (e.g., Wang & Yuan, 2008), which questioned the nature of N400 effect; that is, whether N400 reflects an automatic semantic process or a postlexical process. Regardless, we do not expect to find any delayed N400 in our design since the target stimuli in the current study were always visible and it was the preceding words that were invisible under CFS. Instead of questioning the essence of N400 effect, the reason we focused on N400 component is to use it as an index of semantic integration, which has been well studied (Kutas & Federmeier, 2011). Second, we performed a further analysis by applying the moving average t-test technique across the 0 ms to 1,050 ms time

Table 1 Summary of critical manipulation, measurements, and results in this study

Experiment	Critical manipulation	ERP		Behavior		
		N400		LDT		b-CFS
		Visible	Invisible	Visible	Invisible	Invisible
1a	Prime target congruency (ISI = 400 ~ 600 ms)	+	−			
				−[a]	−	
1b	Statistical power enhanced (double the number of trials)				−	
2a	Attention enhanced (Go-nogo task)		−			
2b	Attention enhanced (LDT task)				−	
3	Same contrast in visible/invisible conditions (ISI = 0 ms)	+	−			
4a	Temporal integration (visible condition first)			+	−	
4b	Spatial integration (visible condition first)			+	−	
5	Sensitivity of current setting tested (idiom, random, inverted)					+

Note: *ERP* event-related potentials, *LDT* lexical decision task, *b-CFS* breaking continuous flash suppression

+: statistically significant; −: statistically insignificant; [a] $p = .056$

window, with a 40-ms time window moving every 10 ms, a wider time window than previously reported (e.g. van Gaal et al., 2014, p. 4). All *p* values were corrected by using the MATLAB "mafdr" function to control the false discovery rate (FDR). The results also showed no significant difference between the congruent and incongruent endings in the time window or region in the invisible conditions (Experiment 1a: all *p*s > .36; Experiment 2a: all *p*s > .54; Experiment 3: all *p*s > .07). These results suggest that late N400 effect was not observed even when we analyzed the overall time window. Last, the N400 differences in the invisible conditions led to very low Bayes factors (i.e., the Bayes factors in the invisible condition were lower than one third), which favored the null hypothesis that there was no N400 effect when the preceding idiom contexts were suppressed from awareness. In summary, the lack of N400 congruency effect in the invisible conditions reflects the inability to integrate subliminal word sequences, rather than the incapability to detect any possible delayed N400 in the invisible condition.

Based on our present results, temporal integration is not as easily processed as well-established spatial integration in the unconscious situation. Previous studies have shown that semantic information such as word (Sklar et al., 2012), number (Bahrami et al., 2010; Van Opstal, de Lange, & Dehaene, 2011), multiple object grouping (Stein, Kaiser, & Peelen, 2015), and scene gist (Mudrik, Breska, Lamy, & Deouell, 2011; but see Moors, Boelens, van Overwalle, & Wagemans, 2016) can be integrated unconsciously. For example, as mentioned above, Sklar et al. (2012) demonstrated that people could read three-word sentences unconsciously by showing that sentences with context violations broke CFS faster than semantically coherent ones (e.g., "I ironed coffee" vs. "I ironed clothes"). Note that these studies all presented

stimuli simultaneously and thus participants could process these stimuli parallelly (cf., Van Opstal et al., 2011). Moreover, even if people needed to integrate information serially across space, the stimuli still preserved on the display and continuously evoked bottom-up signals. Therefore, the coactivated semantic nodes can be linked and integrated without consciousness even though the visual activation evoked by the masked words is too weak to be perceived. On the other hand, the first three words of the idioms in our experiments are distinct arbitrarily meaningful symbols across time. To understand the holistic meaning of the idioms, one needs to extract the semantic information from each word and integrate them into meaningful contexts sequentially. As a consequence, there are no cumulative bottom-up signals in the time domain, and conscious access of successive distinct words is needed.

Our current findings suggest that when subliminal words are presented sequentially, semantic information must be maintained and linked to the following information in order to be successfully integrated. Such information chaining (Sackur & Dehaene, 2009) and the need for working memory support (Baars & Franklin, 2003; Baddeley, 2003; but see Soto, Mäntylä, & Silvanto, 2011, for an exception) are thought to be closely associated with consciousness. Recently, while working memory has been supported to be preserved unconsciously (see Hassin, 2013; Soto et al., 2011; Soto & Silvanto, 2014, for reviews), studies would limit that to low-level properties (e.g., orientation; Soto et al., 2011) or high-level information with preceding visible memory cue (e.g., face Pan, Lin, Zhao, & Soto, 2014). The current study thus provides some insights about to what extent can working memory do without consciousness. On the other hand, according to the modality appropriateness hypothesis (Welch & Warren, 1980), vision is more suitable for spatial information

processing, whereas audition is more appropriate for temporal processing. Perhaps semantic information of words can be spatially but not temporally integrated without consciousness because vision is relatively ineffective at temporal processing, especially with such high-level unconscious information. It is possible that semantic temporal integration in the current study could be done either with the aid of auditory information, such as a beep, or when it is adopted into an auditory version. These possibilities remain for future studies to verify.

In the present study, we established that consciousness plays an important role in semantic temporal integration, which could be explained by the global workspace theory (Baars, 2005; Dehaene & Changeux, 2011; Dehaene et al., 2006; Dehaene, Kerszberg, & Changeux, 1998). As demonstrated throughout our three experiments (Experiment 1, 3 and 4), the Chinese idiom sequences could only be integrated when the words were visible, and the ability failed when words were invisible even with enhanced top-down attention. Specifically, conscious access of the word sequences is important for multiple distinct semantic information to be integrated. The semantic activation evoked by subliminal words decreases rapidly (Greenwald, Draine, & Abrams, 1996; Kiefer & Spitzer, 2000), with reduced intensity and less distributed brain regions than visible words (Dehaene et al., 2001). Although top-down attention is thought to enhance subliminal processing (Dehaene et al., 2006; Van den Bussche et al., 2010), the limited bottom-up activation may not be sufficient to enter the global workspace (Baars, 2005; Dehaene & Changeux, 2011; Dehaene et al., 2006; Dehaene et al., 1998). Theoretically, global workspace contains a network of neurons with long-ranged axons that enable the ignition of activation and allow information to be globally distributed to multiple related brain systems. It might be the case that only visible words can be encoded by sustained activity in the global workspace, which allows the information to be flexibly shared by other related cortical processors, such as working memory and long-term semantic memory systems. Thus, in line with the global workspace theory, our findings suggest that consciousness provides a gateway that enables multiple networks to cooperate and integrate that temporally segregated information into a holistic meaning.

Despite establishing the important role of consciousness in semantic temporal integration in the current study, there remain some issues that call for future studies to explore. First, while we adopted preceding conscious training to enhance unconscious semantic integration in Experiment 4, semantic temporal integration did not occur without consciousness despite some degree of training and a supposedly better familiarity with the task set. However, a recent study (Atas, Faivre, Timmermans, Cleeremans, & Kouider, 2014) found that people are able to learn a sequence of unidentifiable symbols through operant conditioning. This study opened up the possibility that temporal integration of semantic information could also be learned through trial-by-trial feedbacks. Second, based on the modality appropriateness hypothesis as discussed earlier, it is possible that unconscious visual semantic temporal integration could be done with the aid of auditory information or simply occur within auditory modality alone. Finally, Unlike the backward-masking paradigm that the manipulation of unconsciousness is highly dependent on the presentation duration (Pessoa, 2005; Pessoa, Japee, Sturman, & Ungerleider, 2006, the current study adopted the CFS paradigm that can render visual stimuli completely invisible for longer presentation times. Nevertheless, we still cannot observe unconsciously temporal integration of semantic information with the current setting. Combined with previous results with the backward-masking paradigm, our results suggest that the presenting time does not seem to be a critical factor to influence sequential temporal integration. Further studies will be needed to test other factors that may influence semantic integration of multiple words.

In summary, our study has opened up a number of research directions to be explored, including (a) the role of learning in unconscious temporal integration, (b) the modality differences in the ability of semantic integration without awareness, and (c) besides the presenting time of stimuli, testing other factors that may influence sequentially semantic integration in the unconscious condition.

Conclusions

Investigating the limits of unconscious processing is important to clarify the role of consciousness in human perception. In the current study, we found that semantic information of multiple words can be temporally integrated in the visible condition. However, we found no evidence that integration of subliminal word sequences can be accomplished despite enhanced statistical power, enhanced top-down attention, and conscious training. It might be that the activation of subliminal words is too weak to enter the global neuronal workspace (Baars, 2005; Dehaene & Changeux, 2011; Dehaene et al., 2006; Dehaene et al., 1998) to be maintained and spread to related processors, such as working memory and long-term semantic memory systems. Such information broadcast and cross-network cooperation are highly associated with consciousness. In the scope of the current paradigm, our results suggest that consciousness plays a crucial role in temporal integration of semantic information, which might be a limit of unconscious processing.

References

Atas, A., Faivre, N., Timmermans, B., Cleeremans, A., & Kouider, S. (2014). Nonconscious learning from crowded sequences. *Psychological Science, 25*(1), 113–119.

Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research, 150,* 45–53.

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences, 7*(4), 166–172.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science, 21*(2), 224–233.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition, 17*(04), 784–797.

Carrol, G., & Conklin, K. (2015). Cross language lexical priming extends to formulaic units: Evidence from eye-tracking suggests that this idea 'has legs'. *Bilingualism: Language and Cognition, 20*(20), 1–19. doi:10.1017/S1366728915000103

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., & He, S. (2009). Semantic and subword priming during binocular suppression. *Conciousness and Cognition, 18*(2), 375–382.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200–227.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204–211.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America, 95*(24), 14529–14534.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience, 4*(7), 752–758.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5,* 781.

Faivre, N., & Koch, C. (2014a). Inferring the direction of implied motion depends on visual awareness. *Journal of Vision, 14*(4). doi:10.1167/14.4.4

Faivre, N., & Koch, C. (2014a). Temporal structure coding with and without awareness. *Cognition, 131*(3), 404–414.

Faivre, N., Mudrik, L., Schwartz, N., & Koch, C. (2014). Multisensory integration in complete unawareness: Evidence from audiovisual congruency priming. *Psychological Science, 25*(11), 2006–2016.

Gayet, S., Van der Stigchel, S., & Paffen, C. L. (2014). Breaking continuous flash suppression: Competing for consciousness on the presemantic battlefield. *Frontiers in Psychology, 5,* 460.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273*(5282), 1699–1702.

Güldenpenning, I., Koester, D., Kunde, W., Weigelt, M., & Schack, T. (2011). Motor expertise modulates the unconscious processing of human body postures. *Experimental Brain Research, 213*(4), 383–391.

Hassin, R. R. (2013). Yes it can on the functional abilities of the human unconscious. *Perspectives on Psychological Science, 8*(2), 195–207.

Heyman, T., & Moors, P. (2012). Using interocular suppression and EEG to study semantic processing. *Journal of Neuroscience, 32*(5), 1515–1516.

Heyman, T., & Moors, P. (2014). Frequent words do not break continuous flash suppression differently from infrequent or nonexistent words: Implications for semantic processing of words in the absence of awareness. *PLoS ONE, 9*(8), e104719.

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(1), 166.

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science, 18*(4), 349–355.

Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *Journal of Neuroscience, 31*(38), 13535–13545.

Kawakami, N., & Yoshida, F. (2015). Perceiving a story outside of conscious awareness: When we infer narrative attributes from subliminal sequential stimuli. *Consciousness and Cognition, 33,* 53–66.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: Further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research, 13*(1), 27–39.

Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *Neuroreport, 11*(11), 2401–2407.

Kiesel, A., Kunde, W., Pohl, C., Berner, M. P., & Hoffmann, J. (2009). Playing chess unconsciously. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 292.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36*(14), 1.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Reviews of Psychology, 62,* 621.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. New York, NY: Cambridge University Press.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2, Suppl. 2), 467.

Lin, S.-Y., & Yeh, S.-L. (2015). Unconscious grouping of Chinese characters: Evidence from object-based attention. *Language and Linguistics, 16*(4), 517–533.

Liu, Y., Li, P., Shu, H., Zhang, Q., & Chen, L. (2010). Structure and meaning in Chinese: An ERP study of idioms. *Journal of Neurolinguistics, 23*(6), 615–630.

Moors, P., Boelens, D., van Overwalle, J., & Wagemans, J. (2016). Scene integration without awareness: No conclusive evidence for processing scene congruency during continuous flash suppression. *Psychological Science, 27,* 945–956.

Mudrik, L., Breska, A., Lamy, D., & Deouell, L. Y. (2011). Integration without awareness: Expanding the limits of unconscious processing. *Psychological Science, 22*(6), 764–770.

Noel, J.-P., Wallace, M., & Blake, R. (2015). Cognitive neuroscience: Integration of sight and sound outside of awareness? *Current Biology, 25*(4), R157–R159.

Pan, Y., Lin, B., Zhao, Y., & Soto, D. (2014). Working memory biasing of visual perception without awareness. *Attention, Perception, & Psychophysics, 76*(7), 2051–2062.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1), 8–13.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Pessoa, L. (2005). To what extent are emotional visual stimuli processed without attention and awareness? *Current Opinion in Neurobiology, 15*(2), 188–196.

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex, 16*(3), 366–375.

Prioli, S. C., & Kahan, T. A. (2015). Identifying words that emerge into consciousness: Effects of word valence and unconscious previewing. *Consciousness and Cognition, 35,* 88–97.

Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1617.

Sackur, J., & Dehaene, S. (2009). The cognitive architecture for chaining of two mental operations. *Cognition, 111*(2), 187–211.

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences of the United States of America, 109*(48), 19614–19619.

Soto, D., Mäntylä, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology, 21*(22), R912–R913.

Soto, D., & Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends in Cognitive Science, 18*(10), 520–525.

Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking continuous flash suppression: A new measure of unconscious processing during interocular suppression? *Frontiers in Human Neuroscience, 5,* 1–17.

Stein, T., Kaiser, D., & Peelen, M. V. (2015). Interobject grouping facilitates visual awareness. *Journal of Vision, 15*(8), 1–11.

Stein, T., & Sterzer, P. (2014). Unconscious processing under interocular suppression: Getting the right measure. *Frontiers in Psychology, 5,* 387.

Stein, T., Sterzer, P., & Peelen, M. V. (2012). Privileged detection of conspecifics: Evidence from inversion effects during continuous flash suppression. *Cognition, 125,* 64–79.

Tapia, E., & Breitmeyer, B. G. (2011). Visual consciousness revisited magnocellular and parvocellular contributions to conscious and nonconscious vision. *Psychological Science, 22*(7), 934–942.

Tapia, E., Breitmeyer, B. G., & Shooner, C. R. (2010). Role of task-directed attention in nonconscious and conscious response priming by form and color. *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 74.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 1096–1101.

Van den Bussche, E., Hughes, G., Humbeeck, N. V., & Reynvoet, B. (2010). The relation between consciousness and attention: An empirical study using the priming paradigm. *Consciousness and Cognition, 19*(1), 86–97.

van Gaal, S., Naccache, L., Meuwese, J. D., van Loon, A. M., Leighton, A. H., Cohen, L., & Dehaene, S. (2014). Can the meaning of multiple words be integrated unconsciously? *Philosophical Transactions of the Royal Society B, 369*(1641). doi:10.1098/rstb.2013.0212

Van Opstal, F., de Lange, F. P., & Dehaene, S. (2011). Rapid parallel semantic processing of numbers without awareness. *Cognition, 120*(1), 136–147.

Wang, Q., & Yuan, J. (2008). N400 lexicality effect in highly blurred Chinese words: Evidence for automatic processing. *NeuroReport, 19*(2), 173–178.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*(3), 638–667.

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review, 16*(3), 561–566.

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(4), 1069.

Yang, Y.-H., & Yeh, S.-L. (2011). Accessing the meaning of invisible words. *Consciousness and Cognition, 20*(2), 223–233.

Yang, Y.-H., & Yeh, S.-L. (2014). Unmasking the dichoptic mask by sound: Spatial congruency matters. *Experimental Brain Research, 232*(4), 1109–1116. doi:10.1007/s00221-014-3820-5

Yeh, S. L., He, S., & Cavanagh, P. (2012). Semantic priming from crowded words. *Psychological Science, 23*(6), 608–616.

Zhang, H., Yang, Y., Gu, J., & Ji, F. (2013). ERP correlates of compositionality in Chinese idiom comprehension. *Journal of Neurolinguistics, 26*(1), 89–112.

Zhou, S., Zhou, W., & Chen, X. (2004). Spatiotemporal analysis of ERP during Chinese idiom comprehension. *Brain Topography, 17*(1), 27–37.