BRAIN RESEARCH 1118 (2006) 142–154

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Distinguishing source memory and item memory: Brain potentials at encoding and retrieval

Chunyan Guo[a,*], *Li Duan*[a], *Wen Li*[b], *Ken A. Paller*[b,*]

[a]*Department of Psychology, Capital Normal University, PR China*

[b]*Department of Psychology and Institute for Neuroscience, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA*

ARTICLE INFO

Article history:
Accepted 6 August 2006
Available online 14 September 2006

Keywords:
Event-related potential
ERP
Declarative memory
Explicit memory
Recollection
Familiarity

Abbreviations:
ERPs, event-related potentials
Dm, neurophysiological differences at encoding based on later memory performance
ANOVA, analysis of variance

ABSTRACT

Vivid memory for an episode generally includes memory for a central object or event plus memory for background context or source information. To assess neural differences between source and item memory, we used event-related potentials (ERPs) to monitor relevant memory processes at both encoding and retrieval. Participants fluent in Chinese studied Chinese words superimposed on a square or circular background during the study phase, followed by a 1-min delay. Then, memory was tested for both the words (items) and the corresponding background (source), or, in other blocks, tested for the words alone. ERPs to study-phase words differed as a function of whether the word was later remembered. These Dm effects in the interval from 400 to 600 ms, however, did not differ according to whether or not source was remembered. In contrast, ERPs to test-phase words showed clear old/new effects that did differ across conditions. When both item and source were remembered accurately, old/new effects emerged earlier and were larger in amplitude than when source memory was either incorrect or not queried. These results demonstrate that encoding processes indexed by ERPs may have primarily reflected encoding of the visual and semantic properties of these words, stressing item memory over source memory. Retrieval processes indexed by ERPs, in contrast, likely reflected a combination of item retrieval, source retrieval, and related processing engaged when people were remembering words seen earlier.

1. Introduction

When local authorities publicize a photo captured on a surveillance camera of a person suspected of some criminal activity, citizens are supposed to attempt to use that image to probe their memories and answer a critical question—have you seen this face? There are two parts to successful memory in such a circumstance: determining that the face has indeed been seen before and remembering the context of that prior experience. These two parts can be called *item memory* and *source memory*, respectively. Memory for source concerns information beyond the item in question, such as associated sensory features, circumstances of the experience, the time and place, or in short, everything that can add up to make the prior experience a unique episode in one's life. Memory for source is undoubtedly at the core of episodic memory retrieval. A critical issue in memory research, then, is to specify the unique characteristics that differentiate item memory and source memory.

* *Corresponding authors.* C. Guo is to be contacted at Education Science College, Capital Normal University, Beijing, 100037, PR China. Fax: +86 10 68902413. K.A. Paller, Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA. Fax: +1 847 467 1329.

E-mail addresses: guocy@hotmail.com (C. Guo), kap@northwestern.edu (K.A. Paller).

0006-8993/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2006.08.034

The two classic memory experiences of *familiarity* and *recollection* are typically considered to arise on the basis of item and source memory as follows. Veridical memory for an item can allow an individual to recognize that item as familiar, but this experience of familiarity can remain unsubstantiated if the individual cannot also recall source-specifying information (Jacoby and Dallas, 1981; Johnson, 1992; Mandler, 1980). Generally, source memory supplies the particulars to pinpoint a prior episode in which the item was experienced, thus precipitating the full-blown experience of remembering (which is termed recollection). However, perhaps recollection sometimes occurs as the result of strong item memory in the absence of source memory.

In order to fully understand how item memory and source memory differ from each other, both memory encoding and memory retrieval must be considered. There are thus two questions here. First, what brain events at the time of initial encoding mediate memory formation sufficient to later drive accurate item memory versus source memory? Second, what neural differences occur at the time of retrieval when an item is recognized in the absence of source memory compared to when item and source memory occur in tandem?

One way to study memory encoding is to monitor brain electrical activity at the time of initial stimulus processing. For example, event-related potentials (ERPs) elicited by stimuli that were subsequently remembered can be compared to ERPs elicited by stimuli that were subsequently forgotten (Paller and Wagner, 2002). These neurophysiological *d*ifferences based on later *m*emory performance (sometimes referred to generically as *Dm*) provide a way to measure neural events at the time of encoding that are predictive of successful memory performance at some later time. For example, in a few recent studies, memory testing allowed separate Dm analyses according to whether stimuli were recognized on the basis of recollection or familiarity. In a study with facial stimuli, ERP correlates of memory formation differed for faces later recognized with concurrent retrieval of episodic details from the time of learning compared to faces later recognized with familiarity in the absence of episodic recollection (Yovel and Paller, 2004). In a study with pictures of objects, Dm for familiarity exhibited a left frontal topography whereas Dm for recollection was right lateralized for the initial 150 ms and then bilateral (Duarte et al., 2004). In a study with verbal stimuli, young and older subjects showed a Dm for recollection whereas only older subjects showed a Dm for familiarity (Friedman and Trott, 2000b).

Retrieval processing of episodic memory has also been investigated in ERP studies taking into account the contrast between recollection and familiarity (Kounios et al., 2003; Smith, 1993; Trott et al., 1997; Trott et al., 1999; Yovel and Paller, 2004). Many studies of ERPs recorded at retrieval have described effects known as *old/new ERP effects*. These effects refer to the difference between the two ERP waveforms associated with correctly recognized old and correctly rejected new items. In ERP studies of memory retrieval, several kinds of old/new ERP effects related to source memory have been observed (Allan and Rugg, 1998; Cycowicz et al., 2003; Johnson, 1992; Johansson et al., 2002; Kayser et al., 2003; Ranganath and Paller, 1999; Ranganath and Paller, 2000; Van Petten et al., 2000; Wilding, 1999). According to Cycowicz et al. (2001a), one old/new effect occurs at a latency of about 300–700 ms with a posterior scalp distribution, and a second old/new effect has a similar time of onset with a longer duration and a prefrontal scalp distribution. There have been some speculations that item memory is linked with the first ERP effect and source memory with the second (Trott et al., 1999), but such hypotheses are still under active debate.

In addition to these EEG analyses in the time domain, various other methods can also provide relevant evidence concerning encoding and retrieval. For example, in a recent study, magnetoencephalography and frequency-domain analyses were applied to investigate both the encoding and retrieval of declarative memories for outdoor photographs (Osipova et al., 2006). Successful encoding was associated with increased theta power over right temporal regions from approximately 300 to 1000 ms and increased gamma power over the same interval and localized to extrastriate regions. Fairly similar effects were observed at retrieval. Useful evidence can also be obtained when such approaches are applied to study source memory, as in the magnetoencephalographic study of Guderian and Duzel (2005) suggesting that theta oscillations among many different brain regions reflect source retrieval during recollection.

Source memory and item memory have typically been studied in experiments in which participants study a list of items associated with one of two sources. For example, stimuli such as pictures (Duarte et al., 2004) or words (Yonelinas, 1999) can be presented to the left or right of fixation. During memory tests in such paradigms, participants can be required to discriminate between old and new items, and for old items also required to categorize items on the basis of the two initial sources (Batchelder and Riefer, 1990; Bayen et al., 1996; Cycowicz et al., 2001b; Senkfor and Van Petten, 1998). For example, subjects can be given three response options to indicate an old item with source A, an old item with source B, or a new item. Another way to test source memory, known as the sequential response method, is to require a two-choice old/new judgment followed by a two-choice source A/source B decision. In contrast, item memory is typically tested with two response options to endorse items as either old or new.

The major goal of the current experiment was to examine differences between source and item memory via ERPs at both encoding and retrieval. Subjects attempted to memorize Chinese word stimuli, each of which was superimposed on either a square or circular background (Fig. 1). Correct source memory was demonstrated when a subject could remember a word from the study phase and also the background upon which the word appeared. Furthermore, we included an ordinary item recognition test and a source memory test. Source was irrelevant in the former test, hereafter called the *item test*, as there were only two response options ("old"/"new"). In contrast, there were three response options ("old-square"/ "old-circle"/"new") in the latter test, hereafter called the *source test*.

An analysis of ERP data recorded at encoding was aimed at determining whether ERPs would differ reliably as a function of later recognition memory in the two tests, and whether such effects would vary across the two tests. For example, memory processing relevant for later displaying

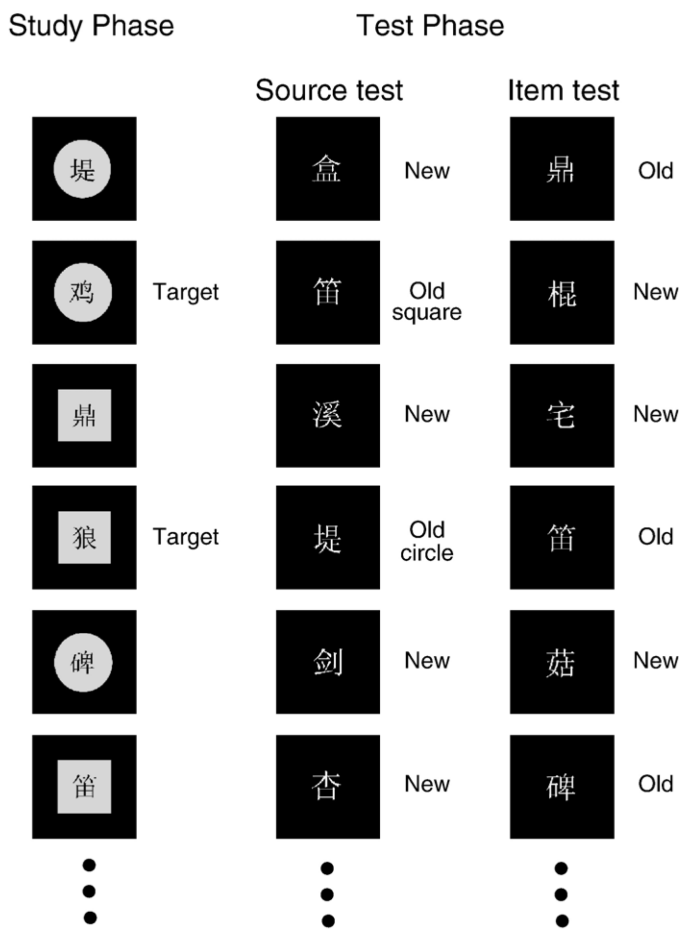

Fig. 1 – Sample stimuli in the two memory tests. Words to the right of stimuli show correct responses that would be made by the subject. In the study phase, subjects attempted to memorize each word while also pressing a button to every target item (animal names). In the test phase, subjects attempted to discriminate old square trials, old circle trials, and new trials (source test) or discriminate old trails and new trials (item test). Retention delay was longer in the item test.

accurate source memory may or may not differ from memory processing relevant for later displaying accurate item memory. Likewise, ERPs were analyzed at retrieval to look for indicators of differences between item retrieval and source retrieval.

2. Results

2.1. *Behavioral results*

In the study phase, accuracy in detecting targets (animal names) was high, with mean rates of 95.2% in the source test and 98.1% in the item test. Demands may have been relatively greater in the source test, given that there was a greater need to encode the background information. In line with this idea, target detection was significantly less accurate in the source test than in the item test [$t(15)=2.70$, $p=0.016$] and also slightly slower in the source test than in the item test [456 ms vs. 422 ms, respectively, $t(15)=2.28$, $p=0.038$].

Results from memory testing are summarized in Table 1. Given the dual response requirements in the source test, we computed two different hit rates (the percentage of old items that received an old response). One hit rate was for correct item and source ($\text{hit}_{\text{item w/ source}}$). The other hit rate was when the source judgment was incorrect ($\text{hit}_{\text{item w/o source}}$). There was only one hit rate in the item test (hit_{item}).

Recognition accuracy was high in both tests. An overall hit rate was computed for the source test as the sum of the $\text{hit}_{\text{item w/ source}}$ and the $\text{hit}_{\text{item w/o source}}$ rates. This combined item hit rate, 68.8%, was significantly higher than the false alarm rate (the percentage of new items that received an old response), 10.1% [$t(15)=27.4$, $p<0.001$]. Likewise, the hit rate in the item test was significantly higher than the false alarm rate [$t(15)=13.2$, $p<0.001$]. The combined hit rate in the source test was only marginally higher from the hit rate in the item test [$t(15)=1.94$, $p=0.071$]. A direct contrast between the two types of hits in the source memory task revealed that $\text{hit}_{\text{item w/ source}}$ occurred more often than $\text{hit}_{\text{item w/o source}}$ [$t(15)=7.721$, $p<0.001$].

Response times for old trials were faster for the item test compared to the source test [$t(15)=10.5$, $p<0.001$], presumably due to the simpler response requirements in the item test. Reaction times also differed between hit and miss trials. In the item test, response times were faster for hit_{item} than for miss trials [$t(15)=2.34$, $p=0.034$]. In the source test, in contrast, misses were faster compared to $\text{hit}_{\text{item w/ source}}$ trials [$t(15)=$

Table 1 – Memory performance measures across source and item tests

Condition	Percent (SE)	Reaction time in ms (SE)
Source test		
$Hit_{item\ w/\ source}$	42.9 (1.6)	1004 (32)
$Hit_{item\ w/o\ source}$	25.9 (1.6)	993 (31)
Miss	30.0 (1.6)	844 (32)
Correct rejection	89.9 (1.8)	786 (27)
Item test		
Hit_{item}	64.2 (2.4)	785 (27)
Miss	35.4 (2.4)	821 (35)
Correct rejection	80.7 (3.2)	794 (30)

Note. For hit conditions, percent denotes percent endorsed as old ($hit_{item\ w/\ source}$ = completely correct; $hit_{item\ w/o\ source}$ = endorsed as old with the wrong source; hit_{item} = endorsed as old in the item test). For correct rejection and miss conditions, percent denotes percent endorsed as new.

9.99, $p<0.001$] and $hit_{item\ w/o\ source}$ trials [$t(15)=9.06$, $p<0.001$]. Presumably, the additional response options in the source test complicated and prolonged the decision process for old trials in the source test. Response times for correct rejections did not differ across tests [$t(15)=0.5$].

2.2. ERP results from study phase

Study-phase ERPs averaged across all 16 subjects from five midline scalp locations (Fpz, Fz, Cz, Pz, Oz) are shown in Fig. 2, with corresponding topographic maps in Fig. 3. Clear differences as a function of subsequent memory were observed at several latencies. For example, beginning about 200 ms after stimulus onset, ERPs were more positive for words that were subsequently remembered compared to those that were forgotten. This sort of ERP difference is termed a *Dm effect* (Paller and Wagner, 2002).

2.2.1. Dm in source test

Two different types of Dm effects were computed for the source test. The first was based on correct source recognition, such that later $hit_{item\ w/\ source}$ trials were compared to later forgotten trials (i.e., later miss), as shown in the left column of Figs. 2 and 3. A repeated-measures analysis of variance (ANOVA) with two factors, subsequent memory (later remembered/forgotten) and electrode location (Fpz/Fz/Cz/Pz/Oz), was conducted on mean amplitude data for three time intervals (220–280, 300–400, and 400–600 ms). In the first two intervals, there were significant main effects of subsequent memory [$F(1,15)=8.6$, $p=0.01$, $F(1,15)=4.9$, $p<0.05$], indicating that the ERPs were more positive for subsequently remembered than forgotten words. At 400–600 ms, there was a main effect of subsequent memory [$F(1,15)=6.9$, $p=0.019$] and an interaction of subsequent memory by location [$F(4,60)=5.6$, $p=0.005$], indicating that the relative positivity for remembered words was larger at anterior locations Fpz ($p=0.045$), Fz ($p=0.001$), and Cz ($p=0.01$). We labeled this the *item w/ source Dm*.

The second type of Dm effect in the source test was based on trials in which words were remembered as old but with the wrong background (i.e., $hit_{item\ w/o\ source}$), as shown in the middle column of Figs. 2 and 3. A repeated-measures ANOVA with two factors, subsequent memory (later remembered/forgotten) and electrode location (Fpz/Fz/Cz/Pz/Oz), was conducted on mean amplitude data for the same three intervals. Again, at 400–600 ms, there was a main effect of subsequent memory [$F(1,15)=8.3$, $p=0.011$] and an interaction of subsequent memory by location [$F(4,60)=4.2$, $p=0.024$], indicating that the relative positivity for remembered words was larger at anterior locations Fpz ($p=0.024$), Fz ($p=0.014$), and Cz ($p=0.008$). We labeled this the *item w/o source Dm*.

2.2.2. Dm in item test

There was one type of Dm analysis in the item test (right column of Figs. 2 and 3). A repeated-measures ANOVA with two factors, subsequent memory (later remembered/forgotten) and electrode location (Fpz/Fz/Cz/Pz/Oz), was conducted on mean amplitude data. At 220–280 ms, an interaction of subsequent memory by location [$F(4,60)=3.4$, $p=0.042$] indicated more positive responses for subsequent hit than subsequent miss only at frontal sites Fz ($p=0.035$) and Cz ($p=0.053$). Likewise at 400–600 ms, an interaction of subsequent memory by location [$F(4,60)=4.0$, $p=0.039$] indicated more positive potentials for subsequent hit than subsequent miss only at the midline frontal site Fz ($p=0.029$). We labeled this the *item Dm*.

2.2.3. Dm across conditions

Additional tests were run to compare Dm across the three conditions described above. First, the measurement in the 400–600 ms interval from the Fz electrode was used, as summarized in Fig. 4A. These three values were not significantly different from each other [$F(2,45)=0.13$, $p=0.88$]. At Fz from 400 to 600 ms, Dm was significant for all three contrasts, and although the magnitude of these differences appeared similar, there could nonetheless be differences across other scalp locations. Dm topography was thus compared across the three conditions following root-mean-square normalization (McCarthy and Wood, 1985), and results failed to show any significant topographic differences [$F(122,1830)=1.90$, $p=0.135$]. Finally, Dm amplitudes were also compared in the two earlier intervals at Fz, and differences across the three conditions were nonsignificant [220–280 ms, $F(2,45)=0.86$, $p=0.43$; 300–400 ms, $F(2,45)=0.25$, $p=0.78$].

2.3. ERP results from test phase

Test-phase ERPs were analyzed separately for each condition. Clear differences were observed beginning about 300 ms after stimulus onset. ERPs were generally more positive for remembered words than for new words. This sort of ERP difference is termed an *old/new effect*.

2.3.1. Old/new effects in source test

Fig. 5 shows ERPs from the three chief conditions in the test phase of the source test. Two different old/new effects were computed. The first was based on correct source recognition, such that $hit_{item\ w/\ source}$ trials were compared to correct rejections. A repeated-measures ANOVA with two factors, condition (old/new) and electrode location (Fpz/Fz/Cz/Pz/Oz), was conducted on mean amplitude data for three time intervals, 300–400, 400–500, and 500–600 ms.

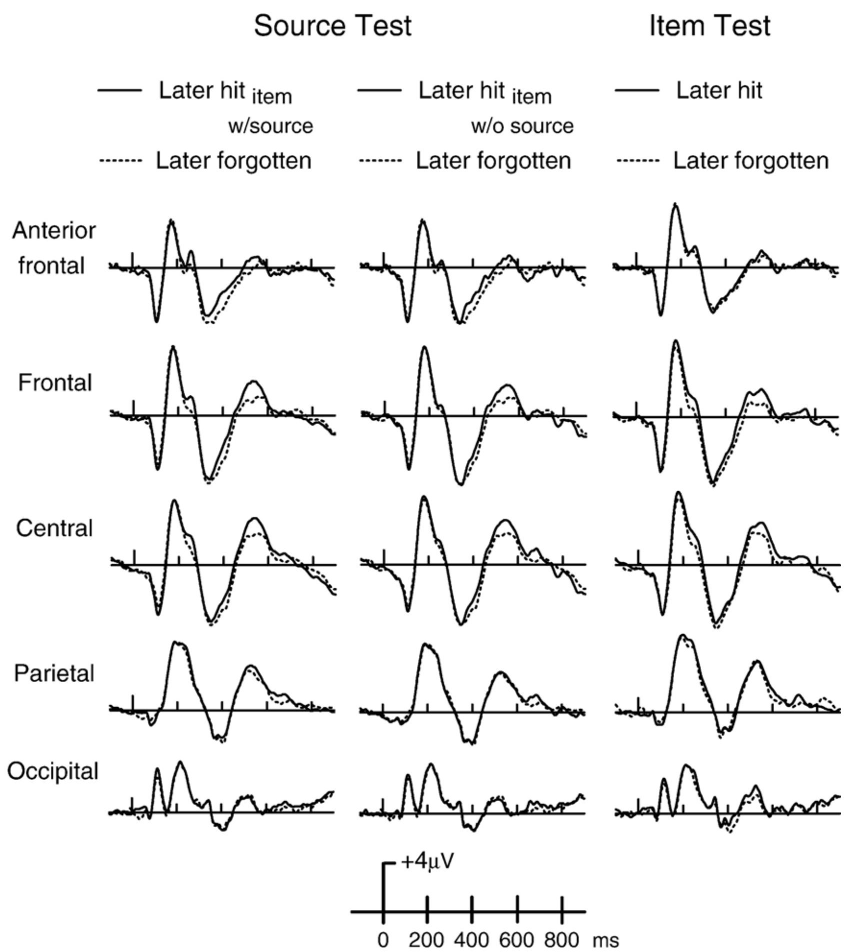

Fig. 2 – ERP results recorded during the study phase from five electrodes along the midline. Superimposed waveforms contrast ERPs for later hit trials with ERPs for later forgotten trials.

A main effect of condition was observed for all three intervals [300–400 ms, $F(1,15) = 10.3$, $p = 0.006$; 400–500 ms, $F(1,15) = 30.1$, $p < 0.001$; 500–600 ms, $F(1,15) = 16.3$, $p = 0.001$]. In all three cases, ERPs were more positive for $hit_{item\ w/\ source}$ compared to correct rejections. We labeled this the *item w/ source old/new effect*.

The second old/new effect was based on recognizing the item as old but with the wrong background; $hit_{item\ w/o\ source}$ trials were compared to correct rejections. For the second two intervals, 400–500 and 500–600 ms, main effects of condition were observed [$F(1,15) = 17.4$, $p = 0.001$, $F(1,15) = 7.9$, $p = 0.013$, respectively]. Again, ERPs were more positive for old than new words. We labeled this the *item w/o source old/new effect*.

Another analysis was conducted to directly compare these two old/new effects. ERP difference waves were computed by subtracting ERPs to correct rejections from ERPs to old words, with separate difference waves for the two types of old words, $hit_{item\ w/\ source}$ and $hit_{item\ w/o\ source}$. Generally, old/new effects were larger in the former case. For example, amplitudes from 500 to 600 ms were larger for $hit_{item\ w/\ source}$ than for $hit_{item\ w/o\ source}$ at Cz [2.8 μV vs 1.9 μV, respectively, $t(15) = 3.2$, $p = 0.006$] and at Fz [2.3 μV vs 1.8 μV, respectively, $t(15) = 2.2$, $p = 0.047$].

Although old/new effect amplitudes differed in these two cases, topographic patterns were similar, as shown in Fig. 5. To directly compare the topography for the two old/new effects, mean amplitudes were normalized using the root-mean-square method (McCarthy and Wood, 1985). The nonsignificant electrode by condition interactions [300–400 ms $F(61,915) = 0.6$; 400–500 ms $F(61,915) = 1.2$; 500–600 ms $F(61,915) = 0.7$] confirmed the lack of evidence for different topographies.

2.3.2. Old/new effects in item test

Fig. 6 shows ERPs from the two chief conditions in the test phase of the item test. ERPs to recognized old words were generally more positive than ERPs to correct rejections. A repeated-measures ANOVA with two factors, condition (old/new) and electrode location (Fpz/Fz/Cz/Pz/Oz), was conducted on mean amplitude data for three time intervals, 300–400, 400–500, and 500–600 ms. For the 400–500 and 500–600 ms intervals, main effects of condition were observed [$F(1,15) = 18.4$, $p = 0.001$, $F(1,15) = 12.7$, $p = 0.003$, respectively]. We labeled this the *item old/new effect*.

An additional analysis was done to directly compare the topography for the old/new effect in the item test with the item w/o source old/new effect in the source test. There was

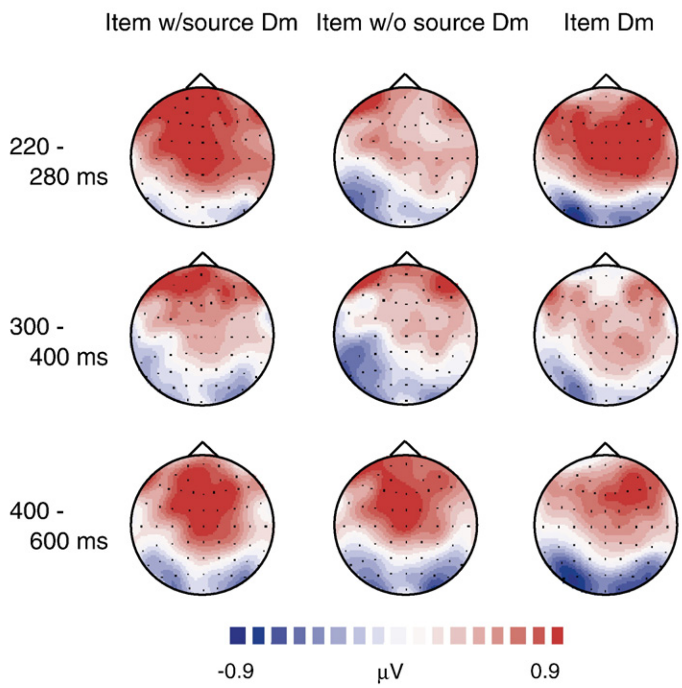

Fig. 3 – Topographic maps in three intervals for differential amplitudes between later hit$_{item\ w/\ source}$ and later forgotten (left column), later hit$_{item\ w/o\ source}$ and later forgotten (middle column), and later hit$_{item\ test}$ and later forgotten (right column).

no evidence for different topographies, as shown by non-significant electrode by condition interactions when normalized mean amplitudes were submitted to ANOVA [300–400 ms $F(61,915)=1.178$; 400–500 ms $F(61,915)=1.6$, $p=0.188$; 500–600 ms $F(61,915)=1.2$].

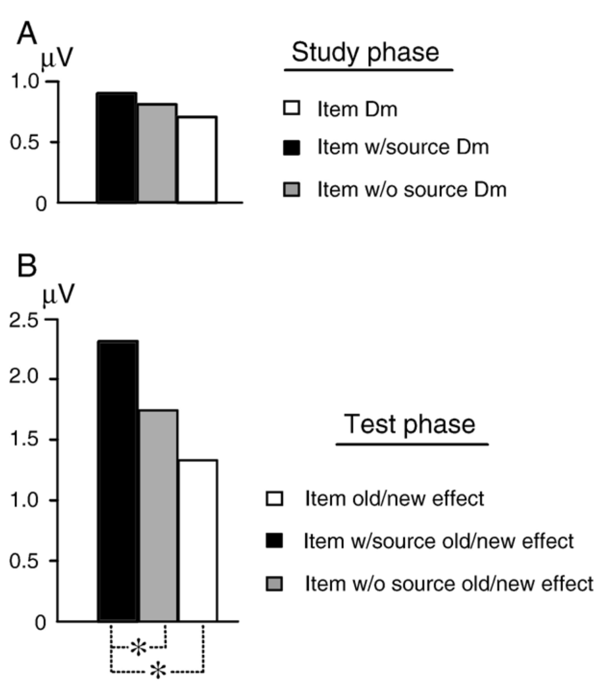

Fig. 4 – Summary of Dm effects in the study phase (A) and old/new effects in the test phase (B), based on mean amplitudes recorded at the frontal midline location (Fz). The 400–600 ms interval was used in A and the 500–600 ms interval was used in B.

2.3.3. *Differences in old/new effects across conditions*

Another test was run to compare old/new effects across the 3 test-phase conditions. The measurement in the 500–600 ms interval from the Fz electrode was used, as summarized in Fig. 4B. Results showed that the item w/ source old/new effect was significantly larger than the item w/o source old/new effect [$t(15)=2.18$, $p<0.05$], and the item w/ source old/new effect was significantly larger than the item old/new effect [$t(15)=2.14$, $p<0.05$]. Both of these contrasts were also significant when amplitudes from all five midline electrodes were used in an ANOVA [$F(1,15)=4.7$, $p<0.05$; $F(1,15)=5.0$, $p<0.05$].

2.3.4. *ERP differences across two tests*

ERPs compared between the two tests are shown in Fig. 7. The earliest apparent difference appeared for both hits and correct rejections at an early latency, about 150–210 ms, when amplitudes were larger in the item test than in the source test. This difference was significant along the midline [$F(1,15)=5.4$, $p=0.034$], and there was also an interaction of test by location [$F(4,60)=10.1$, $p=0.001$] as the effect was only present at anterior locations Fpz ($p=0.004$) and Fz ($p=0.01$). For miss trials, the same sort of effect was shown by a significant task by location interaction [$F(4,60)=9.0$, $p=0.001$], again because the effect was only present at anterior locations Fpz ($p=0.038$) and Fz ($p=0.052$). This effect was unexpected but may reflect the additional processing requirements in the source test. In particular, subjects may use visual imagery during the intertrial interval in considering whether they have remembered the background correctly or not, and this imagery may tax some of the same perceptual networks used for perceptual

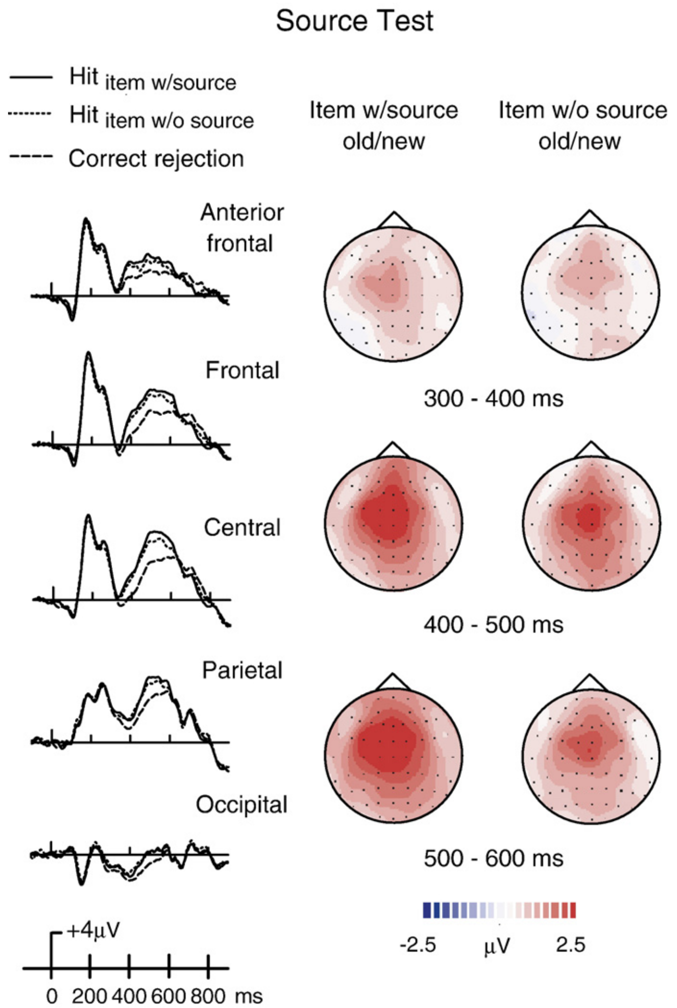

Fig. 5 – ERP waveforms for hit$_{item\ w/\ source}$, hit$_{item\ w/o\ source}$, and correct rejection in the source test (left column); topographic maps for differential amplitudes between hit$_{item\ w/\ source}$ and correct rejection in three time windows (middle column); topographic maps for differential amplitudes between hit$_{item\ w/o\ source}$ and correct rejection in three time windows (right column). (Hit trials are correct responses to old items and correct rejection trials are correct responses to new items.)

analysis, thereby yielding slightly smaller early responses in the source test. Greater task difficulty in the 3-choice source test than in the 2-choice item test is also consistent with the relatively slower reaction times observed in the source test.

Test effects were also analyzed in three intervals from 300 to 600 ms, and no significant task effects were found. This was true both for hits [main effects of task, $Fs \leq 2.2$; interactions $Fs \leq 3.1$] and for correct rejections [main effects of task, $Fs \leq 0.1$; interactions $Fs \leq 1.8$].

2.3.5. Correct rejection versus miss

Old/new effects described above were analyzed using hits and correct rejections, due to the presence and absence of retrieval in these two conditions, respectively. Results were also available for another condition, miss trials, in which retrieval also generally failed to occur. Indeed, inspection of ERPs for correct rejection and miss trials revealed many similarities. The chief latency in which these ERPs differed was from about 500 to 700 ms, when ERPs were relatively more positive for correct rejections. For example, in the source test at midline electrodes, mean amplitudes at 500–700 ms were significantly greater for correct rejections at [Fpz ($p = 0.034$), Fz ($p = 0.003$), Cz ($p = 0.002$), and Pz ($p = 0.002$)]. In the item test at midline electrodes, mean amplitudes at 500–700 ms were significantly greater for correct rejections at Cz ($p = 0.03$). Thus, we infer that old/new ERP effects would appear largely similar if, instead of correct rejections, miss trials were used.

2.4. *Comparison between study-phase and test-phase effects with source localization*

Results form intracranial source analyses calculated using the sLORETA method (see Section 4.5) are shown in Fig. 8. These results correspond to the interval from 400 to 600 ms in the scalp topographic results shown in Figs. 3, 5, and 6). In each case, source analyses conducted showed similar patterns across this time interval; the time points selected for Fig. 8 provide representative findings. The three different Dm effects (Fig. 8A) were similar with respect to the bilateral central–parietal source found with this method. A small inferior

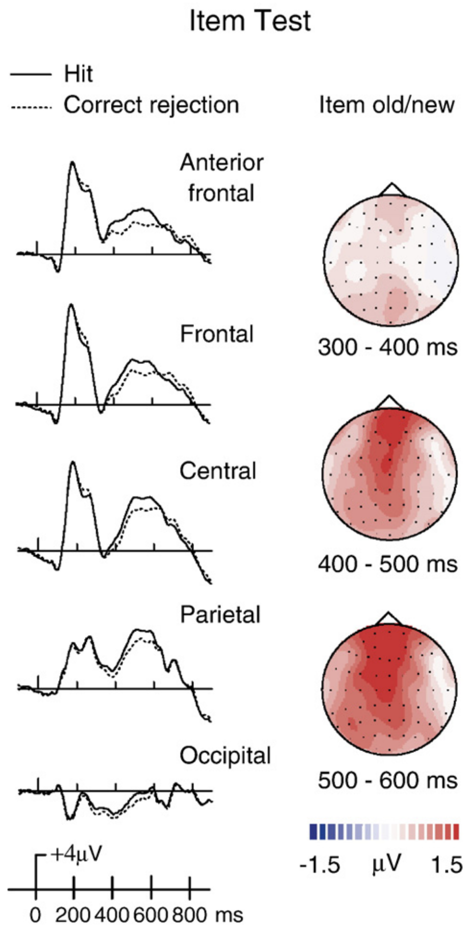

Fig. 6 – ERP waveforms for hit and correct rejection trials in item test (left column); topographic maps for differential amplitudes between hit and correct rejection trials in three time windows (right column).

prefrontal source was also found for the Item Dm, particularly in the right hemisphere, more so than in either of the other two Dm effects. The three kinds of old/new effect (Fig. 8B) also appeared to differ from each other. The item w/ source and item w/o source old/new effects both showed a bilateral central source, whereas the item old/new effect showed a strong bilateral prefrontal source. These differences may reflect different retrieval strategies that were used in the two tests. Interestingly, the prefrontal sources were particularly prominent both at study and test for the item test, whereas central or central–parietal sources were more prominent for the source test. These findings provide interesting clues for direct comparisons between the two tests and between encoding and retrieval (as in the findings of Osipova et al., 2006 that show similarities between sources activated at encoding and at retrieval), but further research is needed to substantiate these intracranial source estimations.

3. Discussion

In the current study, recognition task requirements were manipulated so as to determine whether electrophysiological responses differ between source memory and item memory. To the extent that such differences can be specifically identified, these ERP effects may provide neural clues to the fundamental differences between source memory and item memory, and this information may then help enhance our understanding of the fundamental characteristics of episodic memory in general.

We only manipulated one kind of source information — the background shape upon which words were encoded. Yet, many types of source information can contribute to episodic remembering. For example, Mayes and colleagues have differentiated between extrinsic context, such as spatiotemporal background information, and interactive context, such as spatially adjacent information that can be meaningfully related to target information (Mayes et al., 1985; Mayes et al., 1992). Furthermore, there are many possible sorts of context that vary in the extent to which the information is encoded along with the target (e.g., the location of a word versus the font characteristics versus general environmental factors). Therefore, additional research is required to determine whether the present pattern of findings would also apply for different types of context. Nonetheless, our findings provide some insights into context memory, particularly with respect to the extant literature on ERPs and memory.

Contrasts between source memory and item memory bear an interesting relationship to the distinction between recollection and familiarity, two memory phenomena actively investigated in contemporary memory research (Yonelinas, 2001). When source memory is accurate, a typical inference is

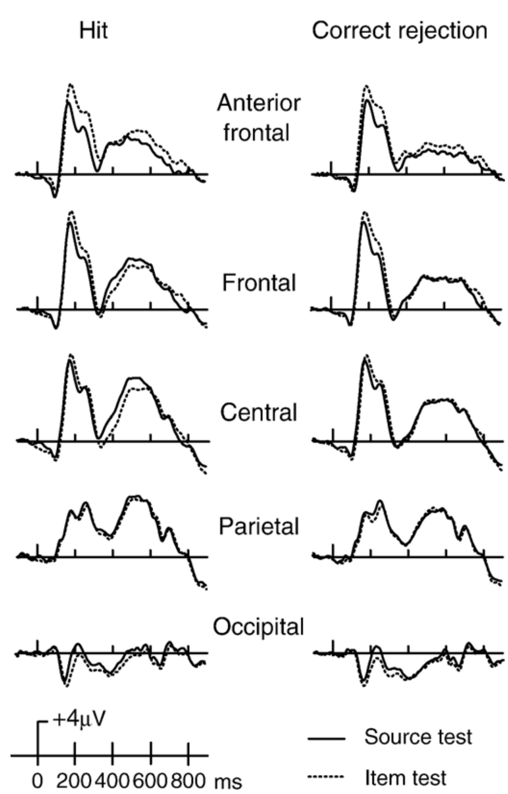

Fig. 7 – ERP waveforms for hit trials in item and source tests (left column) and for correct rejection trials in both tests (right column).

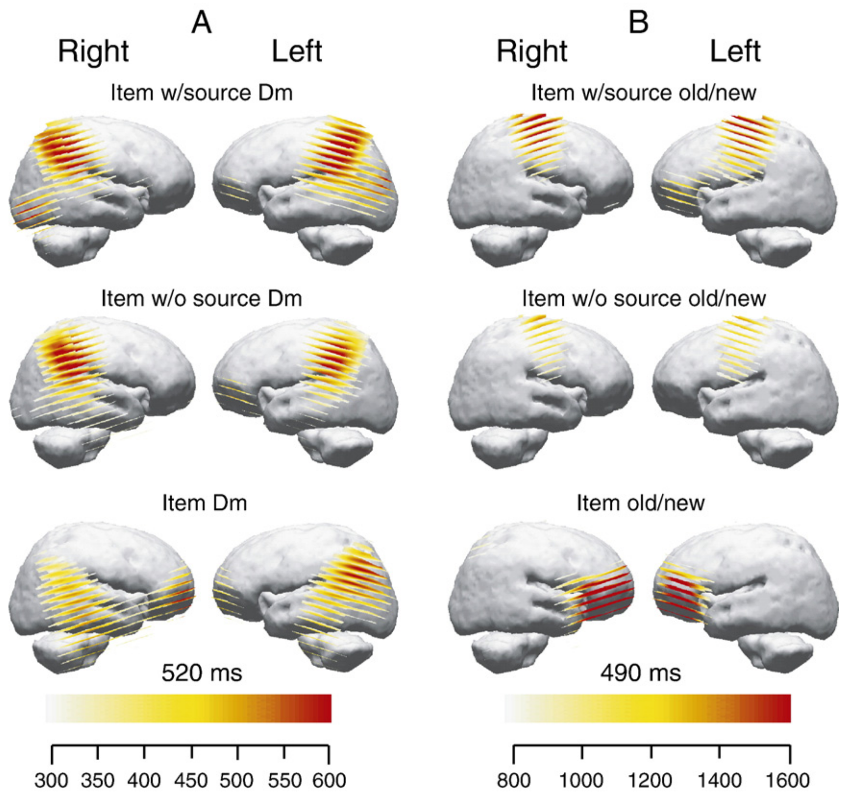

Fig. 8 – Results from intracranial source analysis for ERP differences from the study phase (A) and the test phase (B). Low-resolution current density reconstructions based on the sLORETA model (Pascual-Marqui, 2002) are shown via a color scale for *F* values (current divided by error) as computed at the designated time point superimposed on lateral views of the left and right hemispheres.

that recollection also occurs; in general, when an item is remembered along with contextual information regarding the learning episode, the inference of prior occurrence often follows. In our experiment, contextual information can include many sensory features, one of which was the background shape upon which the word appeared (i.e., the criterion for source memory in this experiment). To the extent that this criterial source information tends to be remembered on trials in which other contextual information is also remembered, criterial source retrieval would be a reasonable indicator of recollection. On the other hand, one cannot infer that correct item retrieval with incorrect source retrieval (i.e., $hit_{item\ w/o\ source}$) indicates that recollection did not occur. A pure familiarity experience (familiarity without recollection) may occur on some of these trials, but recollection may occur on others. In other words, the $hit_{item\ w/o\ source}$ condition cannot be equated with pure familiarity. Nevertheless, there may be some ways in which brain responses on $hit_{item\ w/o\ source}$ trials resemble brain responses that occur in the case of pure familiarity.

Another possible difference between $hit_{item\ w/\ source}$ and $hit_{item\ w/o\ source}$ trials is that item memory may tend to be stronger in the former case. Along with this possible difference, there were likely to have been a higher proportion of guess trials included in the $hit_{item\ w/o\ source}$ condition. The ERP findings, as discussed below, suggest that there were nonetheless qualitative differences between these conditions, consistent with the proposal that source memory occurred more consistently in the $hit_{item\ w/\ source}$ condition than in the $hit_{item\ w/o\ source}$ condition.

Experimental findings pertaining to source memory differed with respect to whether data were collected at encoding or retrieval, as summarized in Figs. 4 and 8. We will thus discuss encoding and retrieval effects in turn.

3.1. *Source encoding*

ERPs recorded during the study phase were predictive of later recognition performance, and these Dm effects differed depending on the circumstances of recognition with respect to source information. Dm effects took the form of more positive waveforms for later remembered than later forgotten trials. These effects add to the ERP literature on Dm, as recently reviewed (Friedman and Johnson, 2000a; Paller and Wagner, 2002; Wagner et al., 1999). The present findings replicate prior results showing that Dm effects can be elicited by Chinese words (Guo et al., 2004). Here, Dm effects were apparent in three time intervals (220–280, 300–400, and 400–600 ms).

In the 400–600 ms interval, Dm was significant for all three of the primary conditions, two from the source test (item w/ source Dm, item w/o source Dm) and one from the item test (item Dm). In all three cases, the topography was similarly biased towards anterior locations (Fig. 3) and similar regions of parietal cortex appeared in the source localization (Fig. 8).

Given that this effect was present for the item w/o source condition, we can infer that it did not pertain to encoding of the background shape. Rather, these three effects all likely pertain to the encoding of the meaning of the word, and meaningful word encoding (i.e., deep semantic processing) is well known to promote later memory retrieval. This conclusion holds even though encoding may have differed between the item test and the source test, in that in the item test, participants may have allocated more attention to word encoding and less to background encoding.

The pattern of findings was different in the early 220–280 ms interval. This Dm effect was apparent in only two of the three comparisons (item w/ source Dm and item Dm), with a similar topography in each case (Fig. 3). This pattern opens the possibility that a process specific to source memory was indexed by this effect. However, the absence of an early item w/o source Dm effect could also be explained by insufficient power in combination with a larger proportion of guess responses that were correct in this condition. Given the early latency of this effect, another possibility is that it reflects the early analysis of the configural meaning of the stimulus. As such, this finding is intriguing in that it occurred earlier than Dm in most prior studies with English words and recognition tests, although early Dm effects are sometimes observed (Guo et al., 2004; Smith, 1993). Given that these characters in the Chinese language are interpreted based on visual features, whereas English words depend on alphabetic rather than pictorial or idiographic information, it is interesting to speculate that the early Dm reflects some aspect of word processing that is not universal across languages, but further studies are needed to test such an idea specifically.

The item w/ source Dm also included significant differences in the interval from 300 to 400 ms, and this effect topographically resembled the 220–280 ms Dm effect. Again, this effect may reflect either some aspect of source encoding or some item encoding that was strongest for trials in the $\text{hit}_{\text{item w/ source}}$ condition. Indeed, a likely possibility is that both the early item w/ source Dm and the early item Dm reflect memory formation with respect to either (a) the analysis of the unique features of the stimulus independent of extrinsic contextual information or (b) the episodic experience that reflects the combination of item encoding and general contextual encoding. Given the early latency of this effect, the former possibility may be more likely, but further tests of this hypothesis are needed.

3.2. *Source retrieval*

In the test phase, successful retrieval was systematically associated with late positive potentials. We categorized these effects as three examples of ERP old/new effects: item w/ source old/new effect, item w/o source old/new effect, and item old/new effect. Generally, relative positivity for old trials was observed from 400 to 600 ms with a frontocentral distribution.

In the source test, the finding that the two types of old/new effect were both evident suggests that source memory is not a necessary requirement. Rather, it seems that both item retrieval and source retrieval may be associated with similar potentials. The largest old/new effect was clearly the item w/ source old/new effect, when episodic retrieval was probably most complete. However, the difference between the two old/new effects in the source test was rather small. Therefore, a likely interpretation is that the major part of the old/new ERP effect either reflected item retrieval or did not require retrieval of the background shape. Given that background shape retrieval was not relevant in the item test, the item old/new effect also does not require this aspect of source retrieval, but it could depend on a combination of item retrieval and episodic retrieval more generally.

These results highlight the more general difficulty of inferring whether ERPs are specific to source memory in the sorts of memory experiments generally conducted in this field. In order to isolate source memory between two conditions, such as the $\text{hit}_{\text{item w/ source}}$ and $\text{hit}_{\text{item w/o source}}$ conditions, it is critical to be confident that item memory is matched between those two. In the present experiment, it is indeed possible that the strength of item memory in these two conditions was not the same. Specifically, it is reasonable to speculate that item memory may have been stronger in the $\text{hit}_{\text{item w/ source}}$ condition. In other words, it would be a mistake to strongly conclude that the strength of item memory for $\text{hit}_{\text{item w/ source}}$ trials was the same as that in $\text{hit}_{\text{item w/o source}}$ trials. At any rate, it is difficult to know for sure without an independent measure of the strength of item memory.

Was an electrophysiological signal specific to source retrieval produced in this experiment and, if so, can it be isolated from other signals that occurred concurrently? The contrast most well suited for this purpose would be between $\text{hit}_{\text{item w/ source}}$ trials and $\text{hit}_{\text{item w/o source}}$ trials. As shown in Fig. 5, this ERP contrast revealed only small differences between these two conditions. In particular, reliable differences were found from 500 to 600 ms at Cz and Fz. These ERP differences are neural correlates of the difference between retrieving a robust episodic memory (including the item integrated with the background) versus item retrieval alone.

3.3. *Conclusions*

The present results add to the literature on Dm and old/new ERP effects in several ways. We observed these effects for stimuli consisting of Chinese characters superimposed on a square or circular background, with memory tested in two different tests. Whereas these results extend ERP results previously produced in other languages, an interesting question that remains to be addressed is how these memory phenomena might be influenced by developmental factors due to the lengthy progression of learning over many years as a person becomes a competent reader of Chinese or other idiographic languages where committing a huge number of symbols to memory is necessary. Two different memory tests were used in this experiment. In the source test, subjects were required to discriminate old words from new words and also indicate the context in which a word had appeared earlier. In the item test, no source judgments were required, only old/new discriminations. Dm effects were apparent based on any manner in which subjects indicated prior occurrence and did not differ greatly across conditions. These study-phase ERP findings indicate that relevant encoding processes that gave rise to ERP signatures at the scalp may primarily reflect encoding of the visual properties of words and the meaning of these words rather than separate aspects of item and source information. If these ERPs specifically reflected formation of memory for source, item w/ source Dm would not have been so similar to item w/o source Dm. Determining whether encoding of source produces a unique ERP correlate may require creating circumstances wherein source is remembered while the item is forgotten, which generally may not occur very often. Test-phase ERPs, on the other hand, differed clearly across conditions. ERPs at retrieval likely reflect a combination of item retrieval, source retrieval, and related processing engaged when people were remembering words seen earlier. Some differences at 500–600 ms may specifically reflect source retrieval. It is plausible to speculate further that recollection was pervasive in the present experiment even in the absence of behavioral indications of source memory (i.e., for hit trials in the item test, when source memory was not tested, as well as for hit trials in the source test when source judgments were incorrect).

4. Experimental procedures

4.1. *Subjects*

Sixteen right-handed students (seven male and nine female) participated in the experiment and received monetary compensation. The mean age was 20.1 years (range = 18–24 years).

4.2. *Stimuli*

A total of 1120 Chinese words were selected, with 560 words in a source memory task and 560 words in an item memory task (means = 19.0 and 17.1 occurrences/million (Beijing Language College 1986), respectively; range = 2–39 occurrences/million). An additional 80 animal names were selected for use in the encoding phase. The 560 words in either task were divided into two sets, studied and unstudied (new), with matched word frequency. Furthermore, the word lists were counterbalanced such that each word appeared equally often as a studied word and as an unstudied word, as well as equally often in the source test and in the item test.

Words in the source memory task were randomly assigned to ten lists, with 28 studied, 28 new, and 4 animal words in each list. Words in the item memory task were randomly assigned to three lists (94 studied words, 94 new words, and 14 animal names in the first list; 93 studied words, 93 new words, and 13 animal names in the second and third lists).

4.3. *Procedure*

Each participant performed two memory tests, the source memory test and the item memory test. Half of the participants performed the source test first, while the other half performed the item test first.

The source test was divided into ten blocks, and the item test was divided into three blocks. Each block comprised an encoding phase, a distraction phase, and a test phase. During the encoding phase, words were presented in black at the center of a monitor for 250 ms separated by an interstimulus interval (ISI) ranging from 1200 to 1600 ms. Each word was presented on a green circle or square background, which was equiprobable. Backgrounds subtended a visual angle of $4.5 \times 4.5°$ (Fig. 1). A fixation cross appeared at the central location during the ISI. Participants were instructed to study each presented word and to press a button when an animal name was shown. Half of the participants used the left hand and the other half used the right hand. At the time of the encoding phase, participants were cognizant of the memory requirements that would come into play during the subsequent test phase. The distraction phase followed the encoding phase, wherein participants were asked to subtract 3 from a 3-digit number shown on the screen for 60 s.

In the test phase of the source test, each block consisted of 28 studied words and 28 new words that were presented in white on a black background for 500 ms with an ISI ranging from 1400 to 1800 ms (Fig. 1). Each participant was instructed to press one of three buttons on the response box to indicate whether the word was presented in the prior encoding phase with a circular background (using the thumb of one hand), with a square background (using the index finger of the same hand), or whether it was new (using the thumb of the other hand). The fingers for responses were balanced across participants. The average delay from when a word appeared in the study phase until when it appeared again in the test phase was 145 s.

In the test phase of the item memory task, each block consisted of 94 or 93 studied and 94 or 93 new words presented for 500 ms with an ISI ranging from 1400 to 1800 ms. Participants judged each word according to whether the word was presented in the prior encoding phase or not (old versus new) by pressing one of two buttons using left or right thumb. Response hand for old and new responses was balanced across participants. The average delay from when a word appeared in the study phase until when it appeared again in the test phase was 345 s.

4.4. *ERP recordings*

Electroencephalographic recordings were obtained from 62 scalp sites using tin electrodes embedded in an elastic cap at locations from the extended International 10–20 System. These electrodes were referenced to the right mastoid during recording and re-referenced to the average of the right and left mastoid offline. Two additional channels were used for monitoring horizontal and vertical electrooculographic (EOG) recordings. Impedance was reduced below 5 KΩ. EEG signals were filtered with a band-pass of 0.05–100 Hz and sampled at a rate of 500 Hz. Each epoch lasted 1000 ms, including 100 ms prior to stimulus onset. Trials with a voltage, relative to the 100-ms baseline, exceeding $\pm 75\ \mu V$ at any electrode were excluded from analysis, as were trials with artifacts in the EOG channels.

4.5. *Data analyses*

ERPs were averaged for study-phase data, when words were presented with one of two backgrounds, and for test phase data, when words were presented with a blank background (Fig. 1). In the source test, test-phase trials for old items were classified as $\text{hit}_{\text{item w/ source}}$ if the correct response was made, as $hit_{item\ w/o\ source}$ if the item was endorsed as an old item but with the incorrect background, and as miss if the item was endorsed as new. In the item test, test-phase trials for old items were either classified as hit_{item} (correct) or miss (incorrect). In both tests, new trials were classified as correct rejections if correctly endorsed as new and as false alarms if incorrectly endorsed as old.

Study-phase analyses were conducted as a function of later test-phase performance, using the same conditions (i.e., three possible outcomes in the source test, two possible outcomes in the item test). Trials were excluded if an incorrect response was made in the study-phase target detection task.

ERPs were quantified by measuring mean amplitudes in three latency intervals (220–280, 300–400, and 400–600 ms for the study phase; 300–400, 400–500, and 500–600 ms for the test phase) relative to the mean amplitude of the prestimulus baseline (−100–0 ms, set to 0 μV in figures). These intervals were selected based on visual inspection of grand-average ERPs, given that similar intervals have been used in prior studies of related ERP phenomena. Although initial analyses focused on five midline locations, topographic analyses confirmed that these midline locations captured the most important effects.

An intracranial source analysis was calculated for each time point between 400 and 600 ms, as well as at the time point with the maximal signal strength as estimated by mean global field power. Analyses were conducted for all difference waves (i.e., Dm effects and old/new effects for each experimental condition). We used the sLORETA method (standardized Low Resolution Electromagnetic Tomography, via Curry V5.0), a modification of the minimum norm least squares approach (L2 norm), which involves dividing each current by the size of its associated error, yielding *F* scores of activation rather than current densities. Prior results have shown that sLORETA produces blurred but accurate localizations of point sources (Pascual-Marqui, 2002). The procedure used a realistic volume conductor model derived using a boundary element method with three layers [skin (10 mm), skull (9 mm), and brain (7 mm), with conductivities of 0.3300, 0.0042, and 0.3300, respectively]. Results presented here (Fig. 8) were representative of the sources shown across the time interval from 400 to 600 ms.

For each dependent variable, an ANOVA with repeated measures was performed. All ANOVAs were two-tailed with level of significance set to $\alpha=0.05$ and supplemented with pairwise comparisons or simple effect comparisons when appropriate. Greenhouse–Geisser corrections were reported when necessary. Midline ERP measurements were evaluated using a condition-by-electrode-location ANOVA for each latency interval. Main effects of electrode location are not reported.

Acknowledgments

This research was supported by the Chinese Ministry of Education grant 20040028001 and National Natural Science Foundation of China (No. 30170322, 30570603) to Chunyan Guo. This research was conducted at the Beijing Key Lab. Wen Li and Ken Paller were supported by US NIH grant NS34639 and NSF grant BCS-0518800.

REFERENCES

Allan, K., Rugg, M.D., 1998. Neural correlates of cued recall with and without retrieval of source memory. NeuroReport 9, 3463–3466.

Batchelder, W.H., Riefer, D.M., 1990. Multinomial processing models of source monitoring. Psychol. Rev. 97, 548–564.

Bayen, U.J., Murnane, K., Erdfelder, E., 1996. Source discrimination, item detection, and multinomial models of source monitoring. J. Exper. Psychol., Learn., Mem., Cogn. 22, 197–215.

Beijing Language College, Xian dai Han yu pin lèu ci dian [Modern Chinese Word Frequency Dictionary], Beijing yu yan xue yuan chu ban she, Beijing, 1986.

Cycowicz, Y.M., Friedman, D., Snodgrass, J.G., 2001a. Remembering the color of objects: an ERP investigation of source memory. Cereb. Cortex 11, 322–334.

Cycowicz, Y.M., Friedman, D., Snodgrass, J.G., Duff, M., 2001b. Recognition and source memory for pictures in children and adults. Neuropsychologia 39, 255–267.

Cycowicz, Y.M., Friedman, D., Duff, M., 2003. Pictures and their colors: what do children remember? J. Cogn. Neurosci. 15, 759–768.

Duarte, A., Ranganath, C., Winward, L., Hayward, D., Knight, R.T., 2004. Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. Cogn. Brain Res. 18, 255–272.

Friedman, D., Johnson Jr., R., 2000a. Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. Microsc. Res. Tech. 51, 6–28.

Friedman, D., Trott, C., 2000b. An event-related potential study of encoding in young and older adults. Neuropsychologia 38, 542–557.

Guderian, S., Duzel, E., 2005. Induced theta oscillations mediate large-scale synchrony with mediotemporal areas during recollection in humans. Hippocampus 15, 901–912.

Guo, C., Zhu, Y., Ding, J., Fan, S., Paller, K.A., 2004. An electrophysiological investigation of memory encoding, depth of processing, and word frequency in humans. Neurosci. Lett. 356, 79–82.

Jacoby, L.L., Dallas, M., 1981. On the relationship between autobiographical memory and perceptual learning. J. Exp. Psychol. Gen. 110, 306–340.

Johnson, M.K., 1992. MEM: mechanisms of recollection. J. Cogn. Neurosci. 4, 268–280.

Johnson, M.K., Kounios, J., Nolde, S.F., 1996. Electrophysiological brain activity and memory source monitoring. NeuroReport 7, 2929–2932.

Johansson, M., Stenberg, G., Lindgren, M., Rosen, I., 2002. Memory for perceived and imagined pictures—An event-related potential study. Neuropsychologia 40, 986–1002.

Kayser, J., Fong, R., Tenke, C.E., Bruder, G.E., 2003. Event-related brain potentials during auditory and visual word recognition memory tasks. Cogn. Brain Res. 16, 11–25.

Kounios, J., Bachman, P., Casasanto, D., Grossman, M., Smith, R.W., Yang, W., 2003. Novel concepts mediate word retrieval from human episodic associative memory: evidence from event-related potentials. Neurosci. Lett. 345, 157–160.

Mandler, G., 1980. Recognizing: the judgment of previous occurrence. Psychol. Rev. 87, 252–271.

Mayes, A.R., Meudell, P.R., Pickering, A., 1985. Is organic amnesia caused by a selective deficit in remembering contextual information? Cortex 21, 167–202.

Mayes, A.R., MacDonald, C., Donlan, L., Pears, J., Meudell, P.R., 1992. Amnesics have a disproportionately severe memory deficit for interactive context. Q. J. Exp. Psychol., A 45, 265–297.

McCarthy, G., Wood, C.C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with

analysis of variance models. Electroencephalogr. Clin. Neurophysiol. 62, 203–208.
Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., Jensen, O., 2006. Theta and gamma oscillations predict encoding and retrieval of declarative memory. J. Neurosci. 26, 7523–7531.
Paller, K.A., Wagner, A.D., 2002. Observing the transformation of experience into memory. Trends Cogn. Sci. 6, 93–102.
Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. 24 (Suppl. D), 5–12.
Ranganath, C., Paller, K.A., 1999. Frontal brain potentials during recognition are modulated by requirements to retrieve perceptual detail. Neuron 22, 605–613.
Ranganath, C., Paller, K.A., 2000. Neural correlates of memory retrieval and evaluation. Brain Res. Cogn. Brain Res. 9, 209–222.
Senkfor, A.J., Van Petten, C., 1998. Who said what? An event-related potential investigation of source and item memory. J. Exper. Psychol., Learn., Mem., Cogn. 24, 1005–1025.
Smith, M.E., 1993. Neurophysiological manifestations of recollective experience during recognition memory judgments. J. Cogn. Neurosci. 5, 1–13.
Trott, C.T., Friedman, D., Ritter, W., Fabiani, M., 1997. Item and source memory: differential age effects revealed by event-related potentials. NeuroReport 8, 3373–3378.
Trott, C.T., Friedman, D., Ritter, W., Fabiani, M., Snodgrass, J.G., 1999. Episodic priming and memory for temporal source: event-related potentials reveal age-related differences in prefrontal functioning. Psychol. Aging 14, 390–413.
Van Petten, C., Senkfor, A.J., Newberg, W.M., 2000. Memory for drawings in locations: spatial source memory and event-related potentials. Psychophysiology 37, 551–564.
Wagner, A.D., Koutstaal, W., Schacter, D.L., 1999. When encoding yields remembering: insights from event-related neuroimaging. Philos. Trans. R. Soc. Lond., B Biol. Sci. 354, 1307–1324.
Wilding, E.L., 1999. Separating retrieval strategies from retrieval success: an event-related potential study of source memory. Neuropsychologia 37, 441–454.
Yonelinas, A.P., 1999. The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. J. Exper. Psychol., Learn., Mem., Cogn. 25, 1415–1434.
Yonelinas, A.P., 2001. Components of episodic memory: the contribution of recollection and familiarity. Philos. Trans. R. Soc. Lond., B Biol. Sci. 356, 1363–1374.
Yovel, G., Paller, K.A., 2004. The neural basis of the butcher-on-the-bus phenomenon: when a face seems familiar but is not remembered. NeuroImage 21, 789–800.