

Research article

Neurophysiological responses during the binding process in working memory

Jacqueline Marcué-Arana, Selene Cansino^{*,1}

Laboratory of NeuroCognition, Faculty of Psychology, National Autonomous University of Mexico, Mexico City, Mexico

ARTICLE INFO

Keywords:

Working memory
Visual binding
Verbal binding
Crossmodal binding
Event-related potentials

ABSTRACT

Working memory is a limited-capacity system responsible for handling and temporarily maintaining information. The multicomponent model of working memory includes the episodic buffer, which encodes, retains, and integrates multimodal information from the visuospatial sketchpad and the phonological loop. Although the model is highly accepted, little research has been conducted to examine the binding process in working memory. This research aimed to examine the neurophysiological similarities and differences among three different types of bindings: verbal-verbal, visual-visual, and verbal-visual. Event-related potentials (ERPs) were recorded in 30 participants while two pairs of stimuli from the different types of bindings were presented followed by a single pair. Participants indicated whether the single pair was equal to one of the previous two pairs, even if the stimulus position was changed, or was not equal to any of them. Compared with crossmodal binding, unimodal binding enhanced the amplitude of the positive slow wave (PSW) during encoding and of the P300 component and PSW during retrieval. These ERPs have been linked to processes such as stimulus classification and association mechanisms. The present study demonstrated that different amounts of resources or underlying processes are required for crossmodal bindings than for unimodal bindings within working memory.

1. Introduction

The binding problem refers to the still-unsolved problem of how the brain is capable of integrating unified and meaningful objects from information that comes separately from different sensory systems [1]. This problem has been of extensive interest because it is crucial to understand how we are able to integrate across time and space attributes, objects, physical and emotional contexts, and actions into a complex unitary experience and, consequently, be aware of that experience [2]. Indeed, consciousness depends on binding because, according to Descartes, the mind, i.e., the person who thinks, must be a unitary entity to be able to experience consciousness of itself [3].

There are different types of bindings, such as integrating the features of an object, integrating the parts of an object, binding an object across time or with its spatial location, or binding objects from different modalities [1]. Moreover, binding is involved in all cognitive processes, such as perception, short-term memory, working memory, long-term memory, and action [2]. The present study focused on unimodal and crossmodal binding in working memory, specifically binding within the

verbal and visual modalities and between both modalities. The present study is closely related to the multicomponent model proposed by Baddeley and Hitch [4] because it encompasses storage subsystems to hold verbal and visual information and an episodic buffer [5] capable of binding the verbal and visual information stored in the subsystems with information from long-term memory into complex episodic representations. However, other models are relevant for our actual understanding of working memory, such as the embedded-process model [6], which emphasizes limited capacity, interaction with long-term memory, and the need for attention to process information in working memory.

The binding of verbal [e.g., 7], visual [e.g., 8], or crossmodal verbal and visual information [e.g., 9] has been examined in several behavioral experiments. These studies commonly employed a secondary task to estimate whether specific processes, such as central executive resources [7], encoding strategies [8], or attention [9], are required for binding. Among these processes, only the disruption of attention interfered with binding. With respect to neuroscience experiments and functional magnetic resonance imaging (fMRI), studies have assessed cross-modality between auditory digits and visual locations [10],

* Correspondence to: Laboratorio de NeuroCognición, Facultad de Psicología, Universidad Nacional Autónoma de México, Avenida Universidad 3004, Building D, Room 212, Colonia Copilco Universidad, Ciudad de México 04510, México.

E-mail address: selene@unam.mx (S. Cansino).

¹ ORCID: 0000-0002-5298-7340.

<https://doi.org/10.1016/j.bbr.2024.115132>

Received 5 April 2024; Received in revised form 27 June 2024; Accepted 28 June 2024

Available online 2 July 2024

0166-4328/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

between visual nonwords and unfamiliar faces [11], between visual digits in an arithmetic task, and between locations in a displacement task [12]. These studies revealed that the processing of two modalities invariably elicited greater activation in the dorsolateral prefrontal cortex during encoding and retrieval. Likewise, activation in the parietal cortex [10,11] and premotor cortex [10] has been reported. The binding of letters and locations was assessed in a magnetoencephalography (MEG) study [13] during the retention period. This study revealed greater theta, beta, and gamma frequency activation in the dorsolateral prefrontal cortex, premotor cortex, and posterior parietal cortex in the binding condition than in the single condition. These results were partially replicated with electroencephalography (EEG) recordings during the retention of letter and location bindings [14]. In this study, the theta band, but not the beta and gamma bands, showed greater frontal power and synchronization between bilateral frontal regions and between frontal and temporal-parietal regions in the binding condition than in the single retention of letters or locations condition.

The indicated MEG and EEG studies have focused on the retention period, whereas event-related potential (ERP) studies have examined neurophysiological brain responses that are directly associated with the encoding and retrieval of stimuli bound. ERPs have been recorded to examine visual binding for two objects, two locations, object and location [15], color and location [16], and color and shape [17]. Compared with similar features, the binding of different visual features elicited lower N100 amplitudes during encoding and lower P200 amplitudes during encoding and retrieval [15]. Both the N100 and P200 have been related to attention. The N2pc component, which is also an indicator of allocation of attention, showed greater amplitude for the binding of different visual features than for single features [16]. P300 exhibited greater amplitude during the retrieval of different feature bindings than during the retrieval of similar feature bindings [15] or a single feature [17]. The P300 has been related to attention and working memory [e.g., 18].

Crossmodal binding examined with ERPs has focused on the binding of visual and semantic information [17,19] and auditory and visual stimuli [20,21]. The P300 showed greater amplitude during retrieval for visual and semantic binding than for two visual features or a single feature [17]. A late slow wave, associated with postretrieval monitoring, showed a lower amplitude for low semantic binding than for high semantic binding to visual stimuli [19]. However, these studies of visual and semantic binding did not actually assess verbal-visual binding, as was done in the current study; specifically, one of the studies involved words being presented as cues during retrieval [17], and the other study involved semantic information that was used as a hint to encode the visual information during encoding [19]. Essentially, participants were not requested to encode and retrieve verbal information. The binding of auditory and visual information elicited greater amplitude at latencies that correspond to the P300 than the sum of the signal for only auditory and visual encoding [21] and the presence of the N2pc and N2ac, which are associated with visual and auditory selective attention, respectively [20].

Previous ERP studies have focused on the binding of similar or different visual features or the crossmodal binding between visual and semantic information and auditory and visual stimuli. These studies demonstrated that the main ERP components that have been recorded during the encoding or retrieval of different types of stimulus binding are the N100, P200, P300, and a positive slow wave (PSW). The N100 is associated with the activation of attention, and the P200 is associated with the allocation of attention [22]. These components indicate the crucial role of attention in the binding process, as demonstrated by pioneering studies on the subject [23]. However, the component observed at 100 ms poststimulus could also manifest as a positive wave, which is identified as the P100. The amplitude of this component has been interpreted as an increase in sensory control mechanisms for attended stimuli [24]. Although the P200 has been mainly related to attention, there is some evidence that the P200 may also index the initial

conscious identification of a stimulus because its amplitude correlates with working memory performance [22]. The P300 amplitude increases when low-probability events occur, thus indicating that it echoes orient attention responses [18]. However, a concept that is more relevant to the present study is the finding that the P300 amplitude is associated with working memory load or difficulty [18]. The PSW waveform has been associated with the engagement of more executive resources and monitoring processes when the retrieval of information is uncertain or more challenging [19]. Given the consistent presence of the N100, the P200, the P300, and the PSW in previous studies and the relevant mechanisms indicated by these waveforms, the present study examined these components to provide more information on the mechanisms underlying the binding process.

As has been reviewed, the binding of verbal and visual information that, according to the multicomponent working memory model [5], relies on the episodic buffer has only been investigated in behavioral, MEG, and EEG studies. Therefore, the aim of the present study was to further contribute to the elucidation of the binding problem by directly contrasting the ERPs elicited by verbal-verbal, visual-visual, and verbal-visual binding. To the best of our knowledge, this is the first ERP experiment that examined these three types of binding in the same study. The following question was addressed in the present study: what are the neurophysiological responses that distinguish the encoding and retrieval of pairs of stimuli of the same modality (verbal-verbal, visual-visual) from pairs of stimuli of different modalities (verbal-visual)? The assessment of verbal-verbal and visual-visual binding would allow us to distinguish the electrophysiological responses that underlie the binding within each subsystem and those that are common to both. The ERPs elicited during crossmodal verbal-visual binding allowed us to examine the electrophysiological responses associated with the episodic buffer and to compare these responses with those related to the unimodal bindings that occur in each subsystem. Moreover, we analyzed ERPs during encoding and retrieval.

To achieve this purpose, we used pairs of letter triads that do not correspond to any meaningful word to ensure that these stimuli would be controlled only by the verbal subsystem without the support of semantic long-term memory. Similarly, we used pairs of abstract forms that were not related to any known object to guarantee that the stimuli would be handled only by the visual subsystem without any verbal support. These same stimuli were combined in pairs to assess verbal-visual binding; thus, the episodic buffer was also examined without the intervention of semantic long-term memory. We used an associative task in which participants had to recognize whether the probe pair was equal to one of the two pairs presented in the study, independent of whether the stimuli within the probe pair had changed their position. This procedure guarantees that the identity of each stimulus within the pair has to be encoded together with its partner to successfully perform the task because merely encoding the pair of stimuli as a rigid visual image or as a sequence of rigid letters is insufficient to perform the task. Moreover, the probe pairs that were not equal to any of the studied pairs were composed of stimuli that belonged to each of the study pairs, i.e., they had already been seen at study; thus, inefficient binding would not allow the detection of these probes because all stimuli would seem familiar.

We anticipate that the two main components associated with attention, the N100 and the P200, will be manifested in the current experiment during encoding and retrieval. These ERP components are expected because the binding process has been linked to attention since the first experiments were performed to elucidate the nature and origin of the binding resources, as outlined above. In particular, we expected that during encoding and retrieval, both components would show greater amplitude during crossmodal binding than during the two unimodal bindings within verbal and visual information. This hypothesis is based on the recurrent finding that the amplitude of both components is modulated by attention [for a review, see 24]. We anticipate that the P300 component will also be expressed during encoding and

retrieval because this component has been consistently recorded during working memory tasks [25]. In particular, we expected that the amplitude of the P300 would be greater during crossmodal binding than during unimodal binding based on the findings [17,21] that crossmodal binding elicited a greater P300 amplitude than the binding of features belonging to the same modality or a single stimulus. We also expected that the PSW would exhibit a lower amplitude during crossmodal binding than during unimodal binding because this waveform exhibits a lower amplitude when more executive and monitoring processes are engaged [19].

2. Materials and methods

2.1. Participants

Thirty healthy right-handed adults participated in the study; half of them were women. The participants were (mean ± standard deviation) 24.21 ± 1.92 years old and had 16.78 ± 1.5 years of formal education. All participants reported no history of psychiatric or neurological disorders, drug or alcohol addiction, or use of central nervous system-altering medication in the past six months. All participants had normal or corrected-to-normal vision, as assessed using the Snellen chart. Participant involvement was voluntary, and each individual provided written informed consent. The study was approved by the Ethics Committee of the Master's and Doctoral Program in Psychology of the National Autonomous University of Mexico in accordance with the principles stated in the Declaration of Helsinki.

2.2. Stimuli

For the verbal-verbal binding task, 120 triads of black letters were employed (e.g., NLT, MQE, and FQY) with no semantic meaning in the Spanish language and that were challenging to pronounce. For the visual-visual binding task, 120 abstract black contour figures that did not represent any object were used; examples of the figures are displayed

in Fig. 1. Triads and figures have the same physical dimensions. The stimuli were randomly selected for each trial. After the whole set was used, the stimuli were again randomly selected for the next trials. Interference from a stimulus that was observed numerous trials ago is unlikely because the visual and verbal stimuli were highly difficult to memorize, which is due to the fact that they had no link with a known object or word. Additionally, the presentation of the stimuli from the subsequent trial is expected to replace the information from the previous trial from working memory due to the limited capacity of working memory. In all binding tasks, the study stimuli consisted of two pairs of triads, two pairs of figures, or two pairs composed of one triad and one figure, according to the binding task: verbal-verbal, visual-visual, and verbal-visual, respectively. During encoding, the two pairs of stimuli were presented simultaneously at the center of the screen, subtending horizontal and vertical visual angles of 4° and 3.5°, respectively. The distance between the two stimuli in the pair was 0.8°. Visual angles were calculated at a distance from the participant's vision and the 'monitor screen of 70 cm.

During retrieval, the probe stimuli consisted of a single pair of triads, figures, or both triads and figures, according to the binding task. The probe could be a pair identical to one of the two pairs previously presented (25 % of the trials), equal to one of the two previous pairs, but the stimuli were presented in different positions (25 % of the trials), or pairs composed of stimuli that were previously presented but in different pairs (50 % of the trials). Thus, all probe stimuli were presented as test stimuli. Two blocks of 60 trials were performed for each binding task. The three types of tasks were presented in counterbalanced order to each participant. The stimuli subtended horizontal and vertical visual angles of 4° and 1.75°, respectively.

2.3. Procedure

Participants attended two sessions. In the first session, general information about the participants was obtained, and visual acuity was evaluated. Then, the participants signed the informed consent letter. During the second session, electrophysiological data were obtained while participants performed the three binding tasks. The second session occurred in a sound-dampened chamber with dim illumination, where participants were seated in a high-back armchair 70 cm away from the monitor. Two response panels with a single button were located above the platform on the right and left sides of the armchair. Participants performed a brief version of the tasks as practice before proceeding to the experimental tasks. The experimental control and collection of behavioral data were performed using EPrime version 2.0 software (Psychology Software Tools, Inc., Pennsylvania, USA).

2.4. Binding tasks

All three binding tasks started with the presentation of a black circle at the center of the screen (with a diameter visual angle of 0.5°) for 200 ms, which served as a fixation point. Following this, the screen remained blank for 200 ms; subsequently, the study stimuli consisted of two pairs of stimuli that were simultaneously presented at the center of the screen for 3000 ms. The pairs were composed of triads, figures, or both triads and figures for the verbal-verbal, visual-visual, and verbal-visual binding tasks, respectively. Afterward, the screen remained blank for 1000 ms, and then the probe stimuli—a pair of triads, figures, or both types according to the task—appeared for 3000 ms. Participants were allowed to respond for a period of 4000 ms after the onset of the probe. Next, a green circle at the center of the screen (with a diameter visual angle of 0.5°) was displayed for 1400 ms to allow participants to blink before the next trial began. Participants were asked to judge whether the single pair of triads, figures, or a combination of both types of stimuli was identical to one of the pairs presented in the study, independent of the position of the stimuli within the pair, or whether the pair was not equal to any of the previous study pairs. Participants

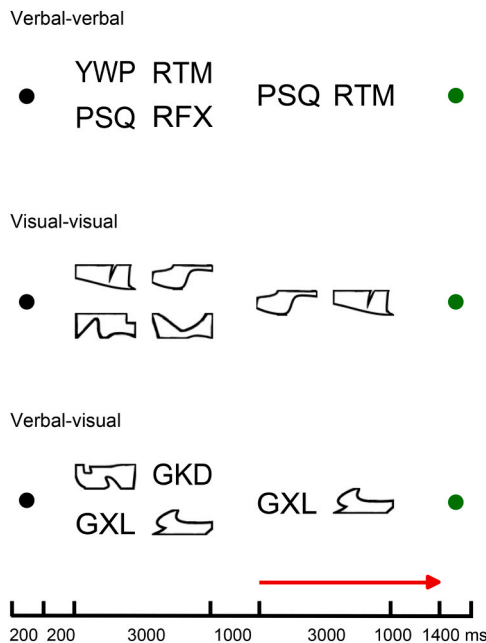


Fig. 1. Sequence of events and their timing in all binding tasks. For the verbal-verbal task, the example represents a trial where the probe pair is different. For the visual-visual task, the example represents a trial where the probe pair is equal but the elements are presented in different positions. For the verbal-visual task, the example represents a trial where the probe pair is equal. The timing of the response is indicated by the red arrow.

provided their responses by pressing one of the keys on the response panels. The keys were counterbalanced and assigned for each response type.

2.5. Event-related potential recording

Electroencephalogram (EEG) recordings were obtained using a BrainVision QuickAmp (Brain Products GmbH; Gilching, Germany) acquisition system with BrainVision Recorder version 1.2 software. Recordings were performed with an elastic ActiCap that included an array of 128 Ag-AgCl active monopolar electrodes. The electrodes were arranged in concentric circles equidistant from the vertex (Cz), which served as an average reference. Vertical electrooculogram (EOG) activity was recorded using two bipolar electrodes placed above and below the left eye. The horizontal EOG was recorded by placing two bipolar electrodes on the outer canthus of each eye. All electrodes had an impedance of less than 10 kΩ. The EEG and EOG signals were continuously digitized at a sampling rate of 500 Hz with a bandwidth between 0.1 and 100 Hz.

2.6. Event-related potential preprocessing

The electrode signals were analyzed by using BrainVision Analyzer version 2.2 software. The preprocessing analyses of the EEG data

initially consisted of offline rereferencing of the original reference (Cz) to TP9 and TP10 (electrodes that correspond to mastoids). Subsequently, ocular correction based on the Gratton et al. [26] algorithm was applied to the signal to minimize the influence of eye movements. Afterward, the signal was digitally filtered offline with a bandwidth of 0.5–20 Hz and a slope of 12 dB/octave. Following this procedure, the EEG signal was segmented into stimulus-locked epochs. ERP epochs lasted 1200 ms, including a 200 ms prestimulus baseline. The duration of the epochs was based on the fact that the main ERPs were manifested in this time period, and no additional waveform variation was noted that could add further information. Epochs with amplitudes that exceeded ± 50 μV or with ocular and motor artifacts were excluded. Additionally, after visual inspection, epochs that were identified with other movement artifacts were manually removed. The epochs were obtained separately during the presentation of the study stimuli and the probe stimuli. Then, the epochs recorded during correct responses for each participant were averaged for each binding task, and the grand averages were obtained. The mean number of epochs was equivalent across all binding tasks [$F(2, 58) = 2.24, p = .128, \epsilon = .80, \eta_p^2 = .07$]: verbal-verbal (mean ± standard error: 95.62 ± 1.94), visual-visual (91.12 ± 2.38), and verbal-visual (94.63 ± 2.11).

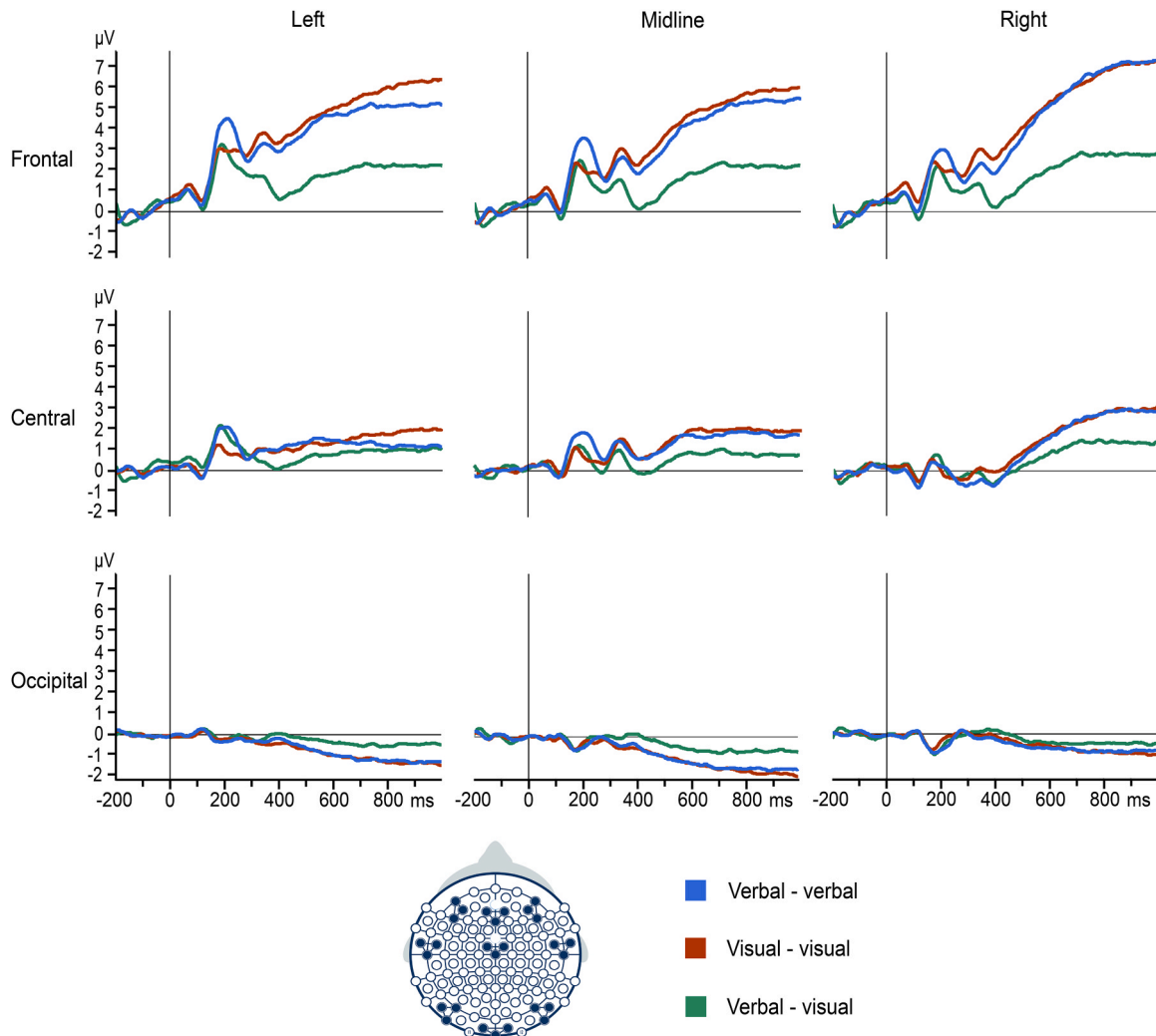


Fig. 2. The grand average ERPs recorded during encoding for each binding task. The insert indicates the electrode sites selected for each region of interest included in the statistical analyses.

2.7. Data analysis

Correct responses and reaction times for correct responses were analyzed separately with repeated-measures ANOVAs. The factor was the binding task (verbal-verbal, visual-visual or verbal-visual). Correct responses for trials containing either the same or different probe pairs were jointly analyzed.

The data from three adjacent electrodes were averaged at nine regions of interest (ROIs): the left frontal, midline frontal, right frontal, left central, midline central, right central, left occipital, midline occipital, and right occipital electrode derivations. Each of these ROIs is depicted in the inserts in Figs. 2 and 3. The ERP mean amplitude was measured relative to the mean voltage during the 200 ms prestimulus baseline over four latency windows during encoding and retrieval (70–150 ms, 150–270 ms, 270–390 ms, and 390–1000 ms). Latency windows were defined according to the timing that was previously reported for measuring each of the ERP components in several studies e.g., [22,25,27,28] and after inspections of the grand average waveforms that were conducted for each binding type. The same latency windows were used to measure the ERP components in all of the binding tasks. For both encoding and retrieval, the first latency window was set to measure a 100-ms component that had a positive amplitude during encoding (P100) and a negative amplitude during retrieval (N100), the second latency window was set to evaluate the P200, the third was set to assess the P300, and the last was set to measure the positive slow wave (PSW).

Repeated measures ANOVAs were conducted separately for each latency window with the factors binding task, region (frontal, central, and occipital), and site (left, midline, and right). Additionally, scalp topography analyses were performed to determine whether the sources that led to each binding modality were different. Scalp topography analyses were performed for the same latency windows used for the ERP analyses. These analyses were conducted on rescaled data using the first procedure proposed by McCarthy and Wood [29] to remove the confounding effects of global differences in magnitude. The raw amplitude was rescaled across binding tasks. Repeated-measures ANOVAs were conducted on the encoding and retrieval data and included binding task and electrode site (128) as factors. Only significant results that involved the interaction of both factors are reported.

To correct for multiple comparisons, we used the Bonferroni procedure. Significant ANOVAs were followed by post hoc comparisons that were computed using the Tukey honest significant difference (HSD) test for within-factor interactions. The effects that did not involve the factor binding task are not described. The degrees of freedom were corrected using the Greenhouse—Geisser procedure to compensate for inhomogeneous covariance that might produce marked increases in type 1 errors. In these cases, the original degrees of freedom, the Greenhouse—Geisser coefficient (ϵ), and corrected probability levels are reported. The significance level was set at $p < 0.05$.

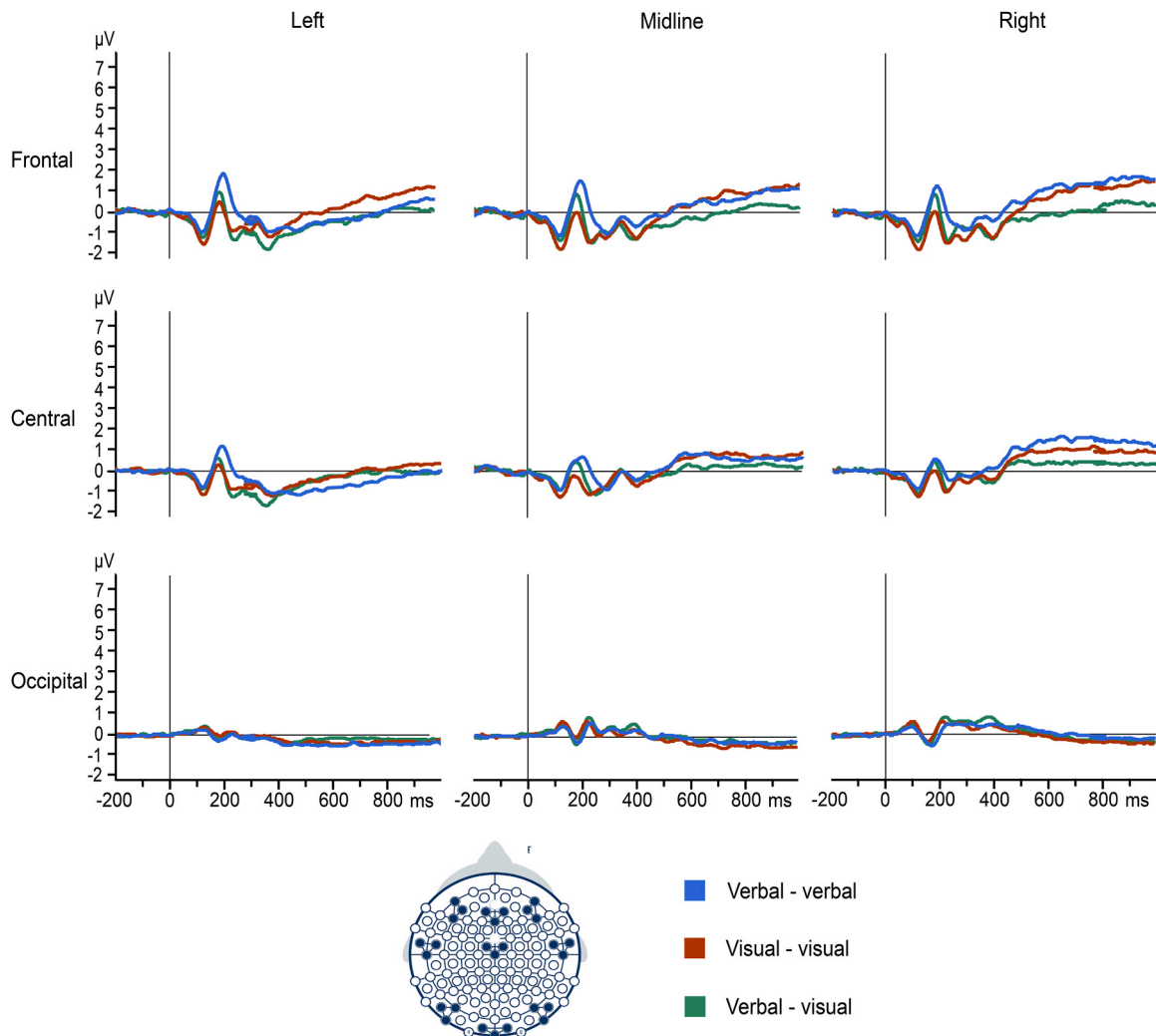


Fig. 3. The grand average ERPs recorded during retrieval for each binding task. The insert indicates the electrode sites selected for each region of interest included in the statistical analyses.

3. Results

3.1. Behavioral results

The repeated-measures ANOVA conducted on correct responses across all types of trials was significant for the factor binding task [$F(2, 58) = 23.09, p < .001, \epsilon = .99, \eta_p^2 = .44$] after applying the appropriate p value cutoff ($p < .016$) according to Bonferroni correction. Tukey HSD post hoc analyses revealed that correct responses were lower in the visual-visual task (78.1 ± 1.5) than in the verbal-verbal task (83.7 ± 1.8) ($p < .001$), which subsequently engendered fewer correct responses than did the verbal-visual task (87.3 ± 1.8) ($p < .001$). The results of the repeated-measures ANOVA conducted on reaction times for correct responses across all types of trials for the factor binding task were also significant [$F(2, 58) = 14.23, p < .001, \epsilon = .87, \eta_p^2 = .33$] after applying the Bonferroni cutoff ($p < .016$). Post hoc analyses revealed that reaction times were longer during the verbal-verbal task ($1452 \text{ ms} \pm 37$) ($p < .001$) and visual-visual task ($1461 \text{ ms} \pm 36$) ($p < .001$) than during the verbal-visual task ($1342 \text{ ms} \pm 36$). No significant difference was noted between the verbal-verbal and visual-visual tasks.

3.2. Event-related potential results

3.2.1. Encoding

The results of the repeated-measures ANOVA conducted during encoding are displayed in Table 1, and the ERPs are depicted in Fig. 2. Fig. 4 displays the amplitude obtained in the three binding tasks in each latency window and ROIs during encoding and retrieval; only the results in which the three-way interaction was significant are shown in this figure. Only results that remained significant after applying the appropriate p value cutoff ($p < .0125$) according to the Bonferroni correction are reported.

3.2.1.1. P200. The three-way interaction between binding task, region and site was significant; post hoc analyses showed that the P200 amplitude during the verbal-verbal task was greater than that observed in the other two tasks across all frontal sites and in the central midline ROI (for both tasks at all ROIs, $p < .001$). Additionally, at the left central ROI, the verbal-verbal ($p < .001$) and verbal-visual ($p < .001$) tasks elicited greater amplitudes than did the visual-visual task.

3.2.1.2. PSW. The analysis of the amplitude of the PSW was significant for the factor binding task. Post hoc analysis revealed that the PSW amplitude was greater for the verbal-verbal (1.96 ± 0.44) ($p = .004$) and

Table 1
ANOVA results on mean amplitude data for all time windows during encoding.

Window (ms)	Factors	F	Df	P	ϵ	η_p^2
70–150	BT	0.45	2,58	.580	.72	.02
	BT x R	1.30	4116	.274	.36	.04
	BT x S	1.95	4116	.156	.46	.06
	BT x R x S	2.46	8232	.037	.61	.08
150–270	BT	2.49	2,58	.109	.74	.08
	BT x R	2.44	4116	.115	.35	.08
	BT x S	3.39	4116	.041	.50	.11
	BT x R x S	3.67	8232	.007*	.52	.11
270–390	BT	2.98	2,58	.078	.71	.09
	BT x R	3.45	4116	.059	.34	.11
	BT x S	1.77	4116	.179	.50	.06
	BT x R x S	2.79	8232	.027	.53	.09
390–1000	BT	9.21	2,58	.001*	.74	.24
	BT x R	9.76	4116	.002*	.33	.25
	BT x S	1.81	4116	.178	.44	.06
	BT x R x S	1.29	8232	.281	.49	.04

Note. BT = binding task (verbal-verbal, visual-visual, verbal-visual), R = region (frontal, central and occipital), S = site (left, midline, right).

* Significant p values after Bonferroni correction.

visual-visual (2.16 ± 0.51) ($p < .001$) tasks than for the verbal-visual task (0.83 ± 0.21). The two-way interaction between the binding task and region was also significant; post hoc tests confirmed that the amplitude of the unimodal tasks (verbal-verbal: $5.26 \pm 1.21, p < .001$; visual-visual: $5.71 \pm 1.41, p < .001$) was greater than that of the verbal-visual task (2.08 ± 0.53) but only at frontal electrode derivations.

3.2.2. Retrieval

3.2.2.1. N100. The results of the repeated-measures ANOVAs conducted on the retrieval data are displayed in Table 2, and the ERPs are shown in Fig. 3. The interaction between the binding task and region factor was significant, and post hoc tests confirmed that there was a difference between the visual-visual (-1.09 ± 0.22) and verbal-verbal (-0.62 ± 0.21) ($p < .001$) tasks but only in the frontal electrode derivations.

3.2.2.2. P200. The results revealed that the factor binding task had a significant effect on the P200; post hoc analysis indicated that the verbal-verbal task (0.24 ± 0.13) elicited a greater amplitude than did the visual-visual (-0.31 ± 0.17) ($p < .001$) and verbal-visual (-0.17 ± 0.14) ($p < .001$) tasks. The interaction between the factors binding task and region was also significant; post hoc results revealed greater amplitudes for the verbal-verbal task than for the other two tasks in the frontal (verbal-verbal: 0.50 ± 0.28 ; visual-visual: $-0.60 \pm 0.33, p < .001$; verbal-visual: $-0.32 \pm 0.29, p < .001$) and central (verbal-verbal: 0.17 ± 0.18 ; visual-visual: $-0.52 \pm 0.24, p < .001$; verbal-visual: $-0.35 \pm 0.21, p < .001$) electrode derivations. The interaction between the binding task and site was also significant; Tukey HSD tests revealed greater amplitudes of the verbal-verbal task than of the other two tasks at the left (verbal-verbal: 0.34 ± 0.17 ; visual-visual: $-0.22 \pm 0.21, p < .001$; verbal-visual: $-0.24 \pm 0.17, p < .001$) and midline (verbal-verbal: 0.25 ± 0.16 ; visual-visual: $-0.40 \pm 0.18, p < .001$; verbal-visual: $-0.17 \pm 0.14, p < .001$) sites. Additionally, the P200 amplitude was greater in the verbal-visual task than in the visual-visual task at midline sites ($p < .001$). Likewise, at the right sites, the verbal-verbal (0.12 ± 0.12) ($p = .029$) and verbal-visual (-0.09 ± 0.15) ($p < .001$) tasks elicited greater amplitudes than did the visual-visual task (-0.31 ± 0.15).

3.2.2.3. P300. The interaction between the factors binding task and site was significant; Tukey HSD analyses showed that the amplitude elicited during the unimodal tasks (verbal-verbal: $-0.33 \pm 0.17, p < .001$; visual-visual: $-0.44 \pm 0.21, p = .003$) was greater than that elicited during the verbal-visual task (-0.77 ± 0.19) at the left sites.

3.2.2.4. PSW. Repeated measures ANOVA of the PSW data was significant for the factor binding task; post hoc analyses showed that the unimodal tasks (verbal-verbal: $0.22 \pm 0.11, p = .005$; visual-visual: $0.21 \pm 0.11, p = .005$) elicited greater amplitudes than did the verbal-visual task (-0.08 ± 0.10). The interaction effect of the factors binding task and region was also significant. Tukey HSD results revealed that in the frontal and central electrode derivations, the amplitude elicited by the unimodal tasks (verbal-verbal: frontal $0.45 \pm 0.23, p < .001$; central $0.34 \pm 0.13, p = .029$; visual-visual: frontal $0.53 \pm 0.24, p < .001$; central $0.34 \pm 0.14, p = .030$) was greater than that elicited during the verbal-visual task (frontal: -0.16 ± 0.22 , central: 0.00 ± 0.14). The interaction between the factors binding task and site was also significant; post hoc analyses showed that the PSW amplitude during the visual-visual task (-0.05 ± 0.14) was greater than that during the other two tasks (verbal-verbal: $-0.37 \pm 0.13, p < .001$; verbal-visual: $-0.31 \pm 0.12, p = .005$) at the left sites. Additionally, at the midline and right sites, the amplitude was greater during the visual-visual task (midline: $0.22 \pm 0.14, p < .001$; right: $0.47 \pm 0.12, p < .001$) than during the verbal-visual task (midline: -0.06 ± 0.12 ; right: 0.12 ± 0.12). The PSW amplitude was also greater during the verbal-verbal task (midline: 0.25 ± 0.13 ; right: 0.78 ± 0.14)

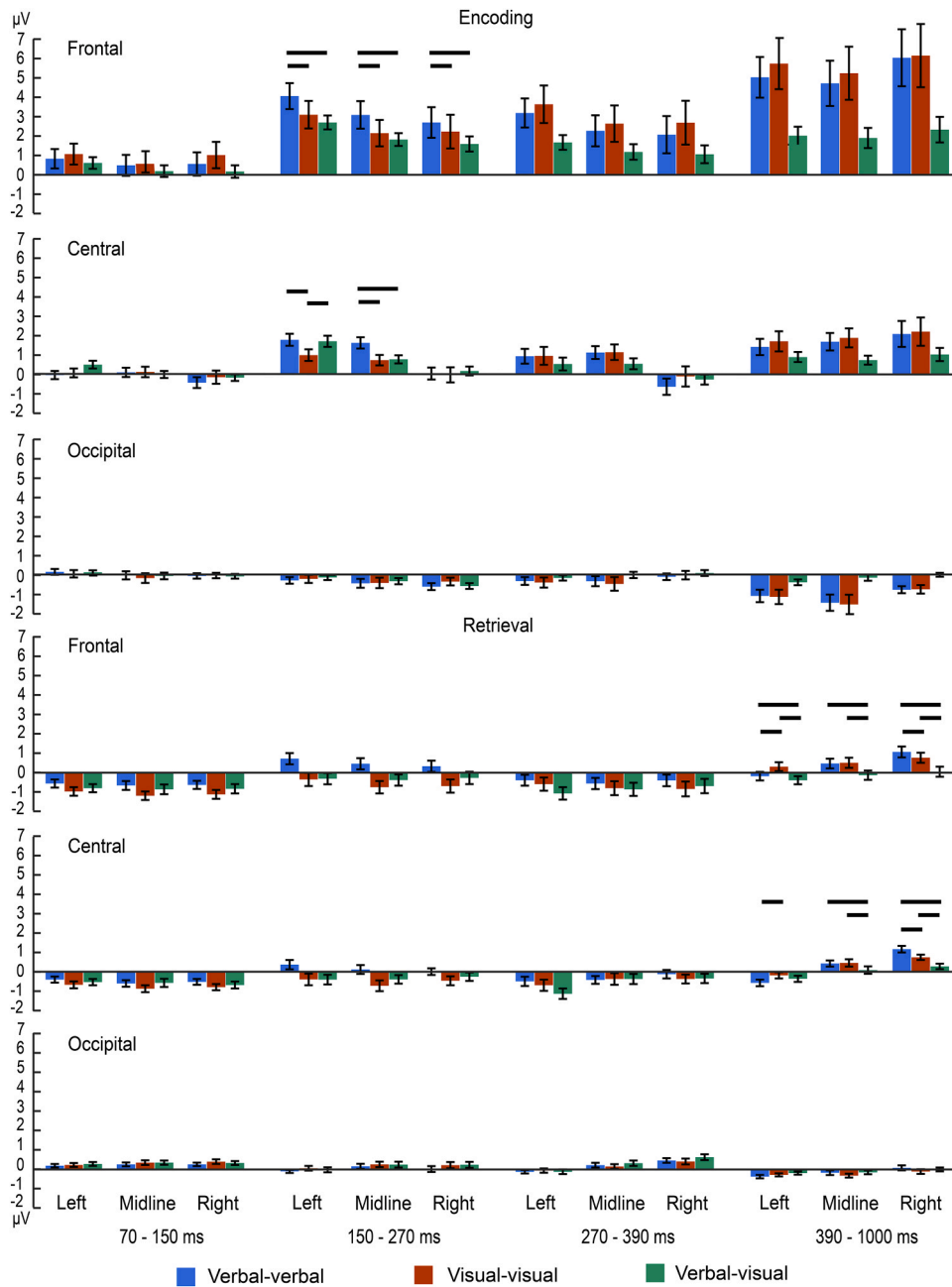


Fig. 4. Mean amplitude during encoding and retrieval at each ROI for each time latency across binding tasks. The horizontal lines indicate significant differences between binding tasks at specific ROIs only when the three-way interaction from the repeated-measures ANOVA was significant. Error bars represent standard errors.

than during the visual-visual task at the right sites ($p < .001$) and during the verbal-visual task at the midline ($p < .001$) and right ($p < .001$) sites. The interaction between the factors binding task, region and site was also significant; post hoc results revealed that the amplitude of the visual-visual task was greater than that of the verbal-verbal task at the left frontal ($p < .001$) and central ROIs ($p < .001$). However, the opposite was observed in the right frontal ($p < .001$) and central ($p < .001$) ROIs. Additionally, the unimodal tasks elicited greater amplitudes than did the verbal-visual task at the frontal midline (verbal-verbal, $p < .001$; visual-visual, $p < .001$), right frontal (verbal-verbal, $p < .001$; visual-visual, $p < .001$), central midline (verbal-verbal, $p < .001$; visual-visual, $p < .001$) and right central (verbal-verbal, $p < .001$; visual-visual, $p < .001$) ROIs.

3.3. Topographic maps

The topographic maps recorded during encoding and retrieval are depicted in Fig. 5. The interaction between binding task and electrode site in the encoding and retrieval data was significant in all latency windows (encoding: 70–150 ms, $F(2, 7366) = 2.15, p < .001, \epsilon = .01, \eta_p^2 = .07$; 150–270 ms $F(2, 7366) = 1.92, p < .001, \epsilon = .01, \eta_p^2 = .06$; 270–390 ms, $F(2, 7366) = 2.84, p < .001, \epsilon = .01, \eta_p^2 = .08$; 390–1000 ms, $F(2, 7366) = 5.15, p < .001, \epsilon = .00, \eta_p^2 = .15$; retrieval: 70–150 ms, $F(2, 7366) = 5.55, p < .001, \epsilon = .01, \eta_p^2 = .16$; 150–270 ms, $F(2, 7366) = 13.42, p < .001, \epsilon = .02, \eta_p^2 = .32$; 270–390 ms, $F(2, 7366) = 1.65, p < .001, \epsilon = .01, \eta_p^2 = .05$; 390–1000 ms, $F(2, 7366) = 6.67, p < .001, \epsilon = .02, \eta_p^2 = .19$). However, post hoc Tukey’s HSD test demonstrated that the scalp topography did not differ among all of the tasks at any electrode site.

Table 2
ANOVA results on mean amplitude data for all time windows during retrieval.

Window (ms)	Factors	F	Df	P	ϵ	η_p^2
70–150	BT	4.71	2,58	.018	.85	.14
	BT x R	5.32	4116	.008*	.50	.16
	BT x S	0.88	4116	.454	.77	.03
	BT x R x S	1.78	8232	.126	.58	.06
150–270	BT	22.21	2,58	<.001*	.99	.43
	BT x R	22.17	4116	<.001*	.52	.43
	BT x S	5.37	4116	.004*	.62	.16
	BT x R x S	1.94	8232	.092	.62	.06
270–390	BT	2.37	2,58	.106	.94	.08
	BT x R	2.59	4116	.083	.51	.08
	BT x S	5.59	4116	.006*	.50	.16
	BT x R x S	2.52	8232	.036	.58	.08
390–1000	BT	7.23	2,58	.002*	.93	.20
	BT x R	9.09	4116	<.001*	.55	.24
	BT x S	17.52	4116	<.001*	.70	.38
	BT x R x S	5.66	8232	<.001*	.66	.16

Note. BT = binding task (verbal-verbal, visual-visual, verbal-visual), R = region (frontal, central and occipital), S = site (left, midline, right).
* Significant p values after Bonferroni correction.

4. Discussion

The main findings of this study were that accuracy and speed were greater in the crossmodal task than in both unimodal tasks. The P200 amplitude was greater for verbal-verbal binding than for the other types of bindings during encoding (frontal ROIs and midline central ROI) and retrieval (frontal and central electrode derivations and at the left and midline sites). The PSW showed lower amplitudes for crossmodal

binding than for unimodal binding during encoding (frontal electrode derivations) and retrieval (frontal, midline and right central ROIs). Likewise, the PSW showed opposite hemispheric amplitude differences between unimodal tasks (visual-visual: left frontal and central ROIs; verbal-verbal: right frontal and central ROIs). Additionally, during retrieval, the N100 component exhibited greater amplitudes (frontal electrode derivations) for visual-visual binding than for verbal-verbal binding, and the P300 component exhibited greater amplitudes for unimodal binding than for crossmodal binding (left electrode sites). Topographic differences were not observed among all types of binding during either encoding or retrieval. Below, we discuss these results in detail.

4.1. Behavioral performance

The crossmodal task was easier to perform than the unimodal tasks, as revealed by the greater number of correct responses and the faster reaction times observed in the verbal-visual task than in the other two tasks. Among the unimodal tasks, the verbal-verbal task was easier to perform than the visual-visual task, as shown by the number of correct responses. The superior performance in the crossmodal task may be attributed to an enhancement of distinctiveness between the members of the pair when they belong to different domains, facilitating their memorability. Another explanation could be that the workload in the crossmodal task for each domain was half that required in the unimodal task, in which four elements of the same domain needed to be retained. Moreover, the performance advantage observed in the crossmodal task provides evidence that the two domains in this task did not compete for the same resources; each of them was processed independently and in

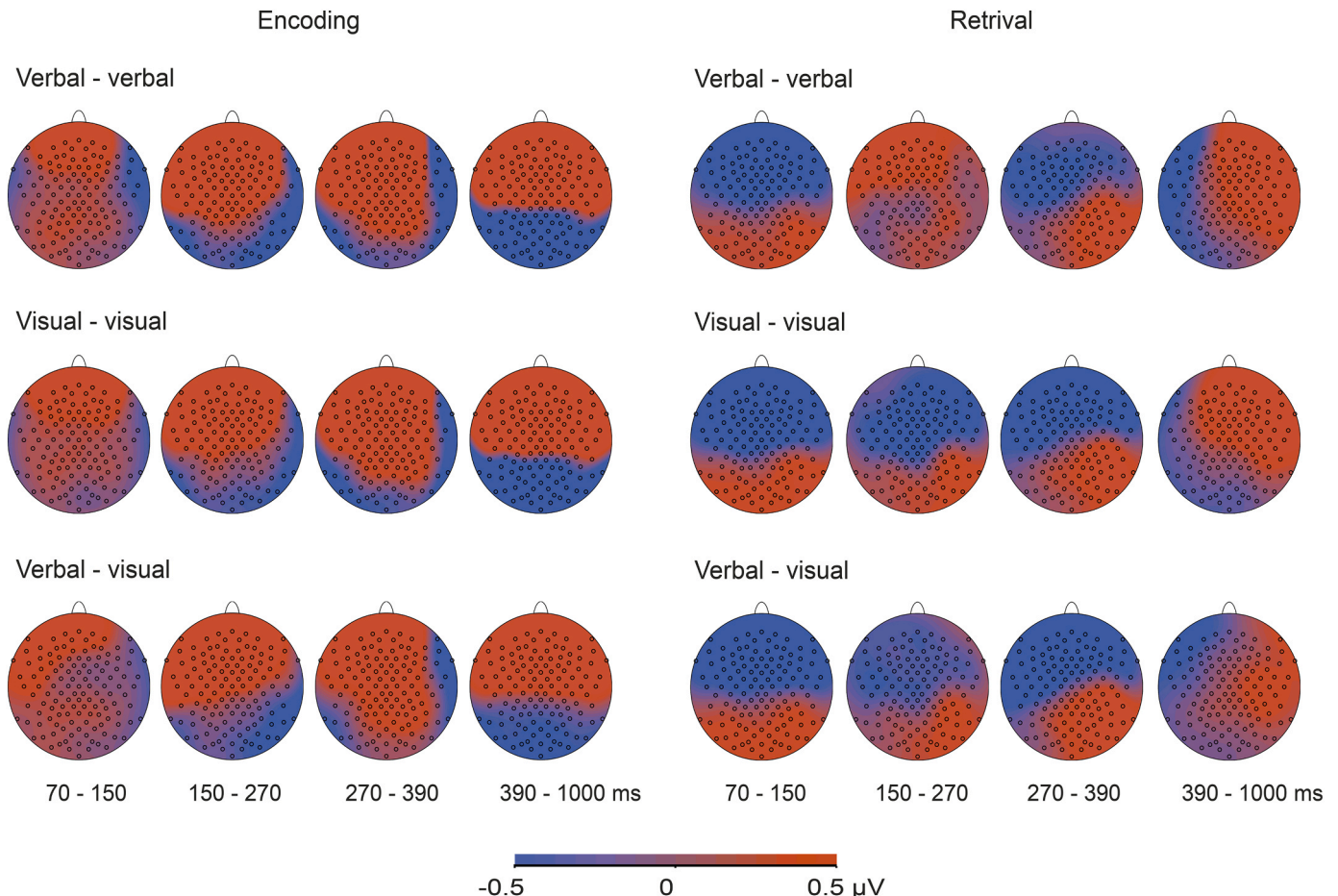


Fig. 5. Topographic maps recorded at each latency window depicting the time course for the three binding tasks during encoding and retrieval.

parallel. Conversely, in both unimodal tasks, the stimuli in each pair were more alike, and the four stimuli were competing for the same resources to be accurately encoded. This effort was even more pronounced in the visual-visual task than in the verbal-verbal task, although both types of stimuli were not related to any semantic information. The recall advantage for triads compared to figures could be related to the fact that individuals are continuously exposed to verbal information, such as letters; therefore, rehearsal abilities are more developed for this type of information. In contrast, retaining figures that have never been seen before may have required novel strategies to encode that were not sufficiently developed to use them accurately.

4.2. Event-related potentials

4.2.1. Encoding

The verbal-verbal task elicited greater P200 amplitudes than did the other two binding tasks at all frontal ROIs, and in the left central ROI, together with the verbal-visual task, it showed greater amplitude than did the visual-visual task. This finding revealed that the P200 amplitude was entirely modulated by verbal information. The P200 component has been related to attention mechanisms and discrimination processes [for a review, see 30]. However, recent studies have shown that the P200 amplitude, recorded at frontal sites, is enhanced during meaningless word learning due to the automatic formation of fast memories [31]. In the present study, the enhanced P200 amplitude for meaningless triads presented along or in combination with figures might have triggered this automatic learning tendency inherent to language.

The PSW amplitude was greater during unimodal tasks than during crossmodal tasks in frontal ROIs. The fact that the PSW was consistently observed in very different experimental procedures led to the interpretation that the PSW was related to the response process, such as response decision, response selection or evaluation of the response provided [32]. However, further research has shed light on the specific process that may be responsible for PSW. According to the current study, PSW is associated with the chunking or binding of information during encoding [33]. In this study, the PSW showed greater amplitude for triads of words semantically related at a frontal derivation. The fact that, in the present study, the PSW was modulated by crossmodal binding provides further evidence that the PSW is a marker of binding that is not limited to verbal information but also to visual information. Moreover, in the present study, we found that the PSW distinguished between unimodal and crossmodal bindings but was present in both modes with different amplitudes. Importantly, this difference should not be attributed to task difficulty because the amplitude of the PSW in the unimodal tasks was equivalent; thus, the PSW did not reflect the difference in task difficulty between the two tasks.

4.2.2. Retrieval

The N100 showed greater amplitude during the visual-visual task than during the verbal-verbal task in the frontal electrode derivations. This component has been associated with the orienting or engaging of attention to a relevant stimulus [34]. In the current study, the greater amplitude of the N100 for the bound visual stimuli compared to that for the bound verbal stimuli may be attributed to the greater difficulty of performing the former task, thus requiring more attentional resources. Although both the P100 and N100 are related to early allocation of attention, the present findings revealed that the N100 devoted attention to internal memory representations.

The P200 component predominated in all tasks that involved verbal information because its amplitude was greater during the verbal-verbal task than during the other two tasks at the frontal and central electrode derivations and at the left and midline sites. Likewise, the verbal-visual task elicited greater P200 amplitudes than did the visual-visual task at midline sites. Together with the verbal-verbal task, its amplitude was also greater than that observed in the visual-visual task at right sites. Therefore, as in encoding, the P200 was highly related to verbal

information. The binding of objects and location was also modulated by the P200 amplitude during encoding and retrieval [15], as observed in the present study for both phases. However, its relevance in the present study for verbal information bounds may be explained by the finding that P200 has been associated with early semantic processes of words, as revealed by a study [35] in which the P200 amplitude was greater for pairs of words semantically related in skilled comprehenders than for those semantically related in less skilled comprehenders. In the present study, the successful retrieval of verbal-verbal bounds may have required verbal abilities that were less necessary for verbal-visual bounds and not necessary for visual-visual bounds.

The P300 component was sensitive to the binding process because its amplitude was greater during unimodal tasks than during crossmodal tasks at left electrode sites. We expected that the crossmodal task would elicit greater P300 amplitudes than the unimodal tasks based on previous studies [17,21], but we observed the opposite. However, the contradictory findings could be clearly explained by the different procedures used among studies. One of them analyzed the binding of visual features [17], and the other compared the binding of visual and auditory information with a single stimulus [21]. In the present study, the P300 showed less amplitude in the crossmodal task, indicating that it was modulated by the binding of information from different domains because its amplitude was equivalent in the unimodal tasks. Moreover, the low P300 amplitude observed in the crossmodal task cannot be attributed to task difficulty because it has been well characterized that the P300 amplitude decreases as workload increases [30]; this was not the case in the present study because behavioral data showed that performance was greater in the crossmodal task than in the unimodal tasks. The P300 component has been related to the categorization of distinct stimuli by means of attention and working memory resources [25], which explains its sensitivity to the binding process that requires such resources.

The unimodal tasks elicited greater PSW amplitudes than did the crossmodal task at midline and right frontal, as well as midline and right central ROIs, indicating that this difference is related to crossmodal bounds. Moreover, this finding revealed that the PSW is a binding marker not only for binding during encoding but also for binding during retrieval. Additionally, the amplitude of the PSW was modulated by the binding of stimuli belonging to a specific domain. The left frontal and central ROIs were more engaged in the binding of figures, whereas the right frontal and central ROIs were more engaged in the binding of triads. Certainly, this explanation is based on the assumption that the manifestation of greater amplitude implies the involvement of more cognitive resources to perform a task. However, this interpretation may be wrong, and further research is still required to unequivocally understand amplitude differences.

4.3. General remarks

One limitation of the present study was that task difficulty was not equivalent among binding tasks. Therefore, task difficulty may have functioned as a confounding variable in the current study. However, we take into account this possibility in all our interpretations. Based on the lack of ERP amplitude correspondence with task difficulty, we determined whether each ERP was related to this variable. Another limitation of this study was that the sample size was not calculated a priori. However, the fact that we found significant results confirmed that the sample size had sufficient power to detect effects. Another limitation of the study is that we did not correlate ERP amplitude with performance or compare correct and incorrect responses for each type of binding. These analyses were not possible because of the low number of incorrect responses; thus, only correct responses were included in the ERP analyses.

The encoding of crossmodal information was modulated by the PSW, and crossmodal information retrieval was modulated by P300 and PSW. All these components showed lower amplitudes during crossmodal

binding than during unimodal binding. Some authors [36] have proposed that a reduction in ERP amplitude may be due to variability in the latency of the components. This latency variability mainly occurs with difficult tasks. However, in the present study, the crossmodal task was less difficult than the unimodal tasks; moreover, no amplitude differences were observed between the unimodal tasks, even though their difficulty was uneven. Additionally, the decrease in amplitude due to latency changes has been associated mainly with the P300 component. In the present study, this decrease was also observed for the PSW component. Therefore, the lower amplitude observed in the crossmodal task compared to that of the unimodal tasks is a genuine finding for the binding of verbal and visual information that reveals its distinctive underlying neural mechanisms. This outcome indicated that the cognitive processes that the P300 and PSW indices were less necessary to bind information from different modalities. Indeed, in the current study, if we were able to contrast only the neurophysiological responses related to the binding process because all tasks requested this function, the lesser amplitude observed for crossmodal binding corresponds to a manifestation of the episodic buffer, which operates as a temporary interface to store the information processed by the subsystems [5]; thus, it is not associated with any additional processing. Further research is required to confirm this outcome.

P300 was initially observed at parietal and central electrode derivations [25]. However, further research has demonstrated that the P300 may be observed at frontal sites, mainly during focal attention and working memory tasks [for a review, see 28]. These findings concur with the left frontal and central distributions observed for the P300 component in the present study. Although we distinguished and measured P300 and PSW as different waveforms that correspond to different underlying processes, there is also the possibility that these two components may be a continuous manifestation of a common process involved in binding. Further research with high-spatial-resolution techniques is required to decipher these two possibilities.

ERP comparisons among the different binding types during encoding were only estimated for correct responses. Therefore, ERPs do not correspond to the so-called subsequent memory effects (SMEs) that predict memorability by comparing stimuli that are later remembered with those stimuli that are later forgotten [37]. Although in the present study, the incorrect responses were not subtracted from the correct responses, the ERPs reflect the neurophysiological responses that were present when efficient encoding led to crossmodal and unimodal binding during retrieval. Independently, from the different analysis approaches, the PSW that was observed during encoding in the present study seems to highly correspond to the late frontal slow wave SME that has been observed in several studies, wherein working memory encoding strategies are engaged to support interitem binding [for a review, see 37].

The ERPs that distinguished between unimodal and crossmodal binding were recorded mainly at the frontal ROIs, followed by the central ROIs during encoding and retrieval. According to fMRI studies [10–12], the processing of two different modalities is associated with greater activation in frontal regions during both encoding and retrieval. The N100 reveals the engagement of attention into memory representations to retrieve bounds. P200 was strongly linked to the binding of verbal information during encoding and retrieval, indicating that this component is related to language learning and preliminary semantic analyses even if triads are meaningless. The PSW recorded at retrieval was the only component that showed a clear hemispheric difference between unimodal tasks. The binding of verbal information was characterized by greater amplitude at the right frontal and central ROIs, whereas the binding of visual information was characterized by greater amplitude at the left frontal and central ROIs. This finding could be related to the engagement of more cognitive resources for each type of binding, but further research is needed to confirm this outcome.

Binding occurs at all levels of cognition and is essential to experiencing consciousness [2]. However, binding does not function equally

within different cognitive processes; therefore, to increase our understanding of binding, it is important to approach its study within specific cognitive processes. In the current study, we were able to examine bindings exclusively within working memory by using abstract forms and nonmeaningful triads of letters that excluded the influence of semantic long-term memory, an influence that has not been excluded in previous studies. Therefore, the current outcomes revealed that working memory function was characterized by engaging more early electrical brain activity to process verbal bindings than other types of bindings and then by assigning an equal amount of brain activity to process both verbal and visual bindings, which was superior to that allocated to crossmodal binding information. This sequence of events occurred during both the encoding and retrieval of information in working memory. However, during retrieval, the earliest electrical activity increase was first attributed to the processing of visual bindings rather than verbal bindings.

Moreover, the present findings provide evidence that the binding process within each of the subsystems outlined by the working memory model may be distinguished by its own neurophysiological pattern. Furthermore, electrophysiological responses were also able to distinguish between unimodal and crossmodal binding, indicating that a different working memory component, probably the so-called episodic buffer, is necessary to unify information from different domains, as revealed by the distinctive neurophysiological manifestations observed when this occurs. The current study contributed to the research on the binding problem within working memory by demonstrating that the binding of items of different modalities is associated with less demanding stimulus classification and associative processes than unimodal binding, as demonstrated by the processes that index the P300 and PSW. Working memory is active during most of our waking time and is essential for daily mental operations that require the crucial ability to bind.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Fundings

This work was supported by the National Autonomous University of Mexico, General Direction of Academic Personal Affairs (DGAPA) (grant number IG300124). J. Marcué-Arana was supported by a PhD scholarship from the National Council for Humanities, Science and Technology (CONAHCYT CVU 960289).

CRediT authorship contribution statement

Selene Cansino: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Jacqueline Marcué-Arana:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization.

Data Availability

Data will be made available on request.

References

- [1] A. Treisman, The binding problem, *Curr. Opin. Neurobiol.* 6 (2) (1996) 171–178, [https://doi.org/10.1016/s0959-4388\(96\)80070-5](https://doi.org/10.1016/s0959-4388(96)80070-5).
- [2] A.L. Roskies, The binding problem, 7–9, *Neuron* 24 (1) (1999) 111–125, [https://doi.org/10.1016/s0896-6273\(00\)80817-x](https://doi.org/10.1016/s0896-6273(00)80817-x).
- [3] A. Brook, P. Raymont, The Unity of Consciousness, in E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition). <https://plato.stanford.edu/>

- archives/sum2021/entries/consciousness-unity/, 2021 (accessed 4 December 2023).
- [4] A.D. Baddeley, G.J. Hitch, Working memory, in: G.A. Bower (Ed.), *Recent Advances in Learning and Motivation*, vol. 8, Academic Press, New York, 1974, pp. 47–89, [https://doi.org/10.1016/s0079-7421\(08\)60452-1](https://doi.org/10.1016/s0079-7421(08)60452-1).
- [5] A.D. Baddeley, The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4 (11) (2000) 417–423, [https://doi.org/10.1016/s1364-6613\(00\)01538-2](https://doi.org/10.1016/s1364-6613(00)01538-2).
- [6] N. Cowan, An embedded-processes model of working memory, in: A. Miyake, P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, Cambridge University Press, Cambridge, 1999, pp. 62–101, <https://doi.org/10.1017/CBO9781139174909.006>.
- [7] A.D. Baddeley, G.J. Hitch, R.J. Allen, Working memory and binding in sentence recall (2009), *J. Mem. Lang.* 61 (3) (2009) 438–456, <https://doi.org/10.1016/j.jml.2009.05.004>.
- [8] A.L. Atkinson, A.D. Baddeley, R.J. Allen, Remember some or remember all? Ageing and strategy effects in visual working memory, *Q. J. Exp. Psychol. (Hove)* 71 (7) (2018) 1561–1573, <https://doi.org/10.1080/17470218.2017.1341537>.
- [9] J.J.R.A. d Santana, C. Galera, Visual-spatial and verbal-spatial binding in working memory, *Psychol. Neurosci.* 7 (3) (2014) 399–406, <https://doi.org/10.3922/j.psns.2014.048>.
- [10] D. Zhang, X. Zhang, X. Sun, Z. Li, Z. Wang, S. He, X. Hu, Cross-modal temporal order memory for auditory digits and visual locations: an fMRI study, *Hum. Brain Mapp.* 22 (4) (2004) 280–289, <https://doi.org/10.1002/hbm.20036>.
- [11] S. Majerus, A. D'Argembeau, T. Martinez Perez, S. Belayachi, M. van der Linden, F. Collette, E. Salmon, R. Seurinck, W. Fias, P. Maquet, The commonality of neural networks for verbal and visual short-term memory, *J. Cogn. Neurosci.* 22 (11) (2010) 2570–2593, <https://doi.org/10.1162/jocn.2009.21378>.
- [12] C. Kim, J.K. Kroger, V.D. Calhoun, V.P. Clark, The role of the frontopolar cortex in manipulation of integrated information in working memory, *Neurosci. Lett.* 595 (2015) 25–29, <https://doi.org/10.1016/j.neulet.2015.03.044>.
- [13] C. Poch, P. Campo, F.B. Parmentier, J.M. Ruiz-Vargas, J.V. Elsley, N.P. Castellanos, F. Maestú, F. del Pozo, Explicit processing of verbal and spatial features during letter-location binding modulates oscillatory activity of a fronto-parietal network, *Neuropsychologia* 48 (13) (2010) 3846–3854, <https://doi.org/10.1016/j.neuropsychologia.2010.09.015>.
- [14] X. Wu, X. Chen, Z. Li, S. Han, D. Zhang, Binding of verbal and spatial information in human working memory involves large-scale neural synchronization at theta frequency, *Neuroimage* 35 (4) (2007) 1654–1662, <https://doi.org/10.1016/j.neuroimage.2007.02.011>.
- [15] E.R. Simon-Thomas, K. Brodsky, C. Willing, R. Sinha, R.T. Knight, Distributed neural activity during object, spatial and integrated processing in humans, *Cogn. Brain. Res.* 16 (3) (2003) 457–467, [https://doi.org/10.1016/s0926-6410\(03\)00060-0](https://doi.org/10.1016/s0926-6410(03)00060-0).
- [16] J.S. Hyun, G.F. Woodman, S.J. Luck, The role of attention in the binding of surface features to locations, *10.1080/13506280802113894*, *Vis. Cogn.* 17 (1–2) (2009), <https://doi.org/10.1080/13506280802113894>.
- [17] Y.F. Niu, B.Z. Shi, L.C. Qiu, B. Li, Q.Y. Shen, C.Q. Xue, Research on binding memory of icon features based on event-related potential, *J. Phys. Conf. Ser.* 1215 (2019) 012027, <https://doi.org/10.1088/1742-6596/1215/1/012027>.
- [18] E. Donchin, Presidential address, 1980. Surprise!... Surprise? *Psychophysiology* 18 (5) (1981) 493–513, <https://doi.org/10.1111/j.1469-8986.1981.tb01815.x>.
- [19] E. Orme, L.A. Brown, L.M. Riby, Retrieval and monitoring processes during visual working memory: an ERP study of the benefit of visual semantics, *Front. Psychol.* 8 (2017) 1080, <https://doi.org/10.3389/fpsyg.2017.01080>.
- [20] L.-I. Klatt, A. Begau, D. Schneider, E. Wascher, S. Getzmann, Cross-modal interactions at the audiovisual cocktail-party revealed by behavior, ERPs, *Neural Oscil. NeuroImage* 271 (2023) 120022, <https://doi.org/10.1016/j.neuroimage.2023.120022>.
- [21] Y. Xie, Y. Xu, C. Bian, M. Li, Semantic congruent audiovisual integration during the encoding stage of working memory: an ERP and sLORETA study, *Sci. Rep.* 7 (2017) 5112, <https://doi.org/10.1038/s41598-017-05471-1>.
- [22] M. Lijffijt, S.D. Lane, S.L. Meier, N.N. Boutros, S. Burroughs, J.L. Steinberg, F. G. Moeller, A.C. Swann, P50, N100, and P200 sensory gating: relationships with behavioral inhibition, attention, and working memory, *Psychophysiology* 46 (5) (2009) 1059–1068, <https://doi.org/10.1111/j.1469-8986.2009.00845.x>.
- [23] A. Treisman, G. Gelade, A feature-integration theory of attention, 197–136, *Cogn. Psychol.* 12 (1) (1980), [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5).
- [24] S.J. Luck, G.F. Woodman, E.K. Vogel, Event-related potential studies of attention, *Trends Cogn. Sci.* 4 (11) (2000) 432–440, [https://doi.org/10.1016/s1364-6613\(00\)01545-x](https://doi.org/10.1016/s1364-6613(00)01545-x).
- [25] A. Kok, On the utility of P3 amplitude as a measure of processing capacity, *Psychophysiology* 38 (3) (2001) 557–577, <https://doi.org/10.1017/s0048577201990559>.
- [26] G. Gratton, M.G. Coles, E. Donchin, A new method for off-line removal of ocular artifact, *Electroencephalogr. Clin. Neurophysiol.* 55 (4) (1983) 468–484, [https://doi.org/10.1016/0013-4694\(83\)90135-9](https://doi.org/10.1016/0013-4694(83)90135-9).
- [27] S.A. Hillyard, L. Anllo-Vento, Event-related brain potentials in the study of visual selective attention, *Proc. Natl. Acad. Sci. USA* 95 (3) (1998) 781–787, <https://doi.org/10.1073/pnas.95.3.781>.
- [28] J. Polich, Updating P300: an integrative theory of P3a and P3b, *Clin. Neurophysiol.* 118 (10) (2007) 2128–2148, <https://doi.org/10.1016/j.clinph.2007.04.019>.
- [29] G. McCarthy, C.C. Wood, Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models, *Electroencephalogr. Clin. Neurophysiol.* 62 (3) (1985) 203–208, [https://doi.org/10.1016/0013-4694\(85\)90858-2](https://doi.org/10.1016/0013-4694(85)90858-2).
- [30] U. Ghani, N. Signal, I.K. Niazi, D. Taylor, ERP based measures of cognitive workload: a review, *Neurosci. Biobehav. Rev.* 118 (2020) 18–26, <https://doi.org/10.1016/j.neubiorev.2020.07.020>.
- [31] B. Bermúdez-Margaretto, Y. Shtyrov, D. Beltrán, F. Cuetos, A. Domínguez, Rapid acquisition of novel written word-forms: ERP evidence, *Behav. Brain. Funct.* 16 (1) (2020) 11, <https://doi.org/10.1186/s12993-020-00173-7>.
- [32] L. García-Larrea, G. Cézanne-Bert, P3, positive slow wave and working memory load: a study on the functional correlates of slow wave activity, *Electroencephalogr. Clin. Neurophysiol.* 108 (3) (1998) 260–273, [https://doi.org/10.1016/s0168-5597\(97\)00085-3](https://doi.org/10.1016/s0168-5597(97)00085-3).
- [33] A.M. Nogueira, O.F. Bueno, G.M. Manzano, A.F. Kohn, S. Pompéia, Late positive slow waves as markers of chunking during encoding, *Front. Psychol.* 6 (2015) 1032, <https://doi.org/10.3389/fpsyg.2015.01032>.
- [34] S.J. Luck, H.J. Heinze, G.R. Mangun, S.A. Hillyard, Visual event-related potentials index focused attention within bilateral stimulus arrays. II. Functional dissociation of P1 and N1 components, *Electroencephalogr. Clin. Neurophysiol.* 75 (6) (1990) 528–542, [https://doi.org/10.1016/0013-4694\(90\)90139-b](https://doi.org/10.1016/0013-4694(90)90139-b).
- [35] N. Landi, C.A. Perfetti, An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders, *Brain Lang.* 102 (1) (2007) 30–45, <https://doi.org/10.1016/j.bandl.2006.11.001>.
- [36] J. Polich, A. Kok, Cognitive and biological determinants of P300: an integrative review, *Biol. Psychol.* 41 (2) (1995) 103–146, [https://doi.org/10.1016/0301-0511\(95\)05130-9](https://doi.org/10.1016/0301-0511(95)05130-9).
- [37] A. Mecklinger, S.M. Kamp, Observing memory encoding while it unfolds: functional interpretation and current debates regarding ERP subsequent memory effects, *Neurosci. Biobehav. Rev.* 153 (2023) 105347, <https://doi.org/10.1016/j.neubiorev.2023.105347>.