Distinct electrophysiological indices of maintenance in auditory and visual short-term memory

Christine Lefebvre a,b,e,*, François Vachon a,b,1, Stephan Grimault a,b,c, Jennifer Thibault a,b, Synthia Guimond a,b,2, Isabelle Peretz a,b, Robert J. Zatorre b,d, Pierre Jolicœur a,b,2,e,*

a Centre de Recherche en Neuropsychologie et Cognition, Département de Psychologie, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, QC, Canada H3C 3J7
b International Laboratory for Brain, Music and Sound Research, Pavillon 1420 Mont-Royal, Université de Montréal, CP 6128, succursale Centre Ville, Montréal, QC, Canada H3C 3J7
c Centre National de Recherche Scientifique, 31 Chemin Joseph Aiguier, 13402 Marseille, Cedex 20, France
d Montreal Neurological Institute, McGill University, 3801 University Street, Montréal, QC, Canada H3A 2B4
e Centre de recherche de l'Institut universitaire de gériatrie de Montréal, 4545 Chemin Queen Mary, Montréal, QC, Canada H3W 1W4

Article info
Article history:
Received 31 October 2012
Received in revised form
30 June 2013
Accepted 2 August 2013
Available online 11 August 2013

Keywords:
Auditory short-term memory
Visual short-term memory
ERP
Sustained anterior negativity
Sustained posterior contralateral negativity

Abstract
We compared the electrophysiological correlates for the maintenance of non-musical tones sequences in auditory short-term memory (ASTM) to those for the short-term maintenance of sequences of coloured disks held in visual short-term memory (VSTM). The visual stimuli yielded a sustained posterior contralateral negativity (SPCN), suggesting that the maintenance of sequences of coloured stimuli engaged structures similar to those involved in the maintenance of simultaneous visual displays. On the other hand, maintenance of acoustic sequences produced a sustained negativity at fronto-central sites. This component is named the Sustained Anterior Negativity (SAN). The amplitude of the SAN increased with increasing load in ASTM and predicted individual differences in the performance. There was no SAN in a control condition with the same auditory stimuli but no memory task, nor one associated with visual memory. These results suggest that the SAN is an index of brain activity related to the maintenance of representations in ASTM that is distinct from the maintenance of representations in VSTM.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Short-term memory (STM) is at the epicentre of the most cognitive activity and as such, it has been studied extensively. STM has been defined, re-defined, measured, used as a diagnostic tool, and theorized about. Much is known about it, but our perspective on STM is somewhat biased because most studies and theories of STM focus on verbal material, whether heard, read, or grasped from an image or a sound. This component is named the Sustained Anterior Negativity (SAN). The amplitude of the SAN increased with increasing load in ASTM and predicted individual differences in the performance. There was no SAN in a control condition with the same auditory stimuli but no memory task, nor one associated with visual memory. These results suggest that the SAN is an index of brain activity related to the maintenance of representations in ASTM that is distinct from the maintenance of representations in VSTM. The study of verbal STM is not only valid, but it is also extremely useful. Humans manipulate and use verbal material very frequently and efficiently, and understanding how the brain manipulates and processes verbal material in STM is thus important. It is also an intuitive choice: we use it overtly and extensively every day, and it is easy to incorporate in experiments. Several models of STM were therefore developed with mostly verbal material in mind. For example, arguably the most influential model of STM, Baddeley’s working memory model (Baddeley, 2000; Baddeley & Hitch, 1974), is composed of a central executive that controls different systems: the visuospatial sketch pad, responsible for the temporary storage and manipulation of spatial and visual information; the phonological loop, for maintenance of material readily recordable in a verbal form; and the episodic buffer, responsible for the storing of multi-dimensional code. When confronted to non-verbal auditory material, one would be tempted to make provision for it within the structures of the model. For example, it would seem that a melody could be maintained in the phonological loop. However, there is evidence that verbal and non-verbal sensory material might be processed

*Corresponding authors at: Université de Montréal, Département de Psychologie, C.P. 6128, succursale Centre-ville, Montréal, Québec, Canada H3C 3J7.
1 Author Vachon is now at École de psychologie, Pavillon Félix-Antoine-Savard, 2325, rue des Bibliothèques, Université Laval, Québec, QC, Canada G1V 0A6.
2 Author Guimond is now at Douglas Mental Health University Institute, McGill University, 3801 University Street, Montréal, QC, Canada H3A 2B4.
independently in STM. For example, it has been shown that pitch is processed independently, in STM, from spoken words and digits (Deutsch, 1970; Rowe & Rowe, 1976; Rowe, Philippchalk, & Cake, 1974), timbre (Semal & Demany, 1991, 1993), and loudness (Clement, Demany, & Semal, 1999; Semal & Demany, 1993). As an example, Deutsch (1970) found that hearing spoken digits did not interfere with the maintenance of the pitch of tones, whereas tones did not interfere with the maintenance of spoken digits. This is evidence that different structures might be used to maintain different types of representations for these different stimuli. It means, in turn, that what we know about verbal STM might not be readily applicable to non-verbal, auditory stimulation, hence the need to study non-verbal, auditory STM (ASTM) on its own. Williamson, Baddeley, and Hitch (2010) also came to such a conclusion after they found that the short-term maintenance of verbal and that of purely acoustic material share few characteristics. The authors suggested that the two types of information are stored in separate stores, namely a phonological loop for verbal material and a tonal loop for acoustic material.

One problem with the suggestion of different maintenance modules for phonological and tonal representations, as pointed out by Postle (2006), is that one may need to postulate a distinct store for timbre (Semal & Demany, 1991, 1993) and possibly stores for all manner of stimuli that might be processed independently in STM. As a result, Postle, as well as D’Esposito (2007), suggest an alternative to the multiple stores assumed in Baddeley’s model. Acknowledging the literature on the role of the prefrontal cortex (PFC) in STM (see Fuster, 2001, for a review), Postle suggests that instead of housing STM stores, the PFC’s role might be one of control and activation of the relevant information in STM. Perhaps the PFC would stimulate activation in the areas responsible for the perception of the relevant stimuli, substituting bottom-up activation with top-down activation. In short, according to Postle, the PFC controls which system remains activated for further processing or maintenance in STM after the original stimuli are gone. Thus, tonal information, for example, would be maintained in structures responsible for the perception of tonal information.

Recent literature on visual STM (VSTM) provides support for this view. For instance, Vogel and Machizawa (2004) recorded electrical brain activity that is closely linked to maintenance of visual material in STM. This activity was recorded at posterior electrode sites, making it unlikely that the PFC was involved. The evidence provided by Vogel and Machizawa was convincing because not only did activation increase with memory load, but it also ceased to increase when maximum capacity was reached despite an increase in the number of presented items. The levelling off of activation when memory capacity was reached suggests a more direct reflection of memory rather than perception, given that the activation followed the functional characteristics of visual memory rather than of visual perception. Moreover, they found a correlation between individual participants’ capacity and the increase in activation with load: the higher the capacity, the larger the increase in activation recorded. Perron et al. (2009) showed further that brain activation related to memory maintenance returned to zero earlier when the retention interval of the memory task was shorter. This pattern of results provided converging evidence for the view that the Sustained Posterior Contralateral Negativity (SPCN, referred to as the CNSW by Klaver, Talsma, Wijers, Heinez, & Mulder, 1999, and the CDA by Vogel & Machizawa, 2004) is an index of the maintenance of visual stimuli in STM. Using fMRI, Todd and Marois (2004) showed that a particular portion of the IPS/IOS (intraparietal sulcus, intraoccipital sulcus) exhibited a pattern of increase in the BOLD (blood oxygenation level-dependent) response that followed a similar pattern as that showed by the SPCN (namely an initial increase in

response followed by a levelling off when memory capacity was reached). Their results suggested that maintenance activity for visual stimuli involves bilateral cortical regions in posterior areas, rather than in the PFC. These results were later further corroborated by magnetoencephalography (MEG) evidence for the involvement of IFS/IOS during the maintenance phase of visual short-term memory experiments (see Robitaille, Grimault, & Jolicoeur, 2009; Robitaille et al., 2010).

It is plausible, therefore, that ASTM is mediated by distinct structures from those subserving VSTM, and those responsible for the maintenance of material that can be retained in a verbal form. Our assumption is that processes relying on different neuronal structures have a good chance of yielding different electrophysiological patterns. Few papers have looked into the neurophysiological underpinnings of ASTM, and fewer have tried to differentiate the structures responsible for the maintenance of non-verbal material within the two modalities. In this paper, we investigated the maintenance of tone sequences in ASTM and compared the associated electric brain activity with that associated with the maintenance of sequences of visual stimuli, using similar experimental designs, enabling direct comparisons of brain activity for auditory and visual memory.

1.1. Auditory short-term memory

We define ASTM as the retention of acoustic properties of perceived sounds and of sequences of perceived sounds for periods of a few seconds. The information stored in ASTM comprises low-level characteristics such as pitch, timbre, and amplitude, devoid of particular meaning other than the sensory information itself. ASTM is not verbal or semantic in nature, and does not refer to phonological content. Because of the fundamental importance of ASTM for cognitive processing from acoustic input, we wish to understand the basic functional and neuronal representation of the storage mechanisms that make ASTM possible using the event-related potential (ERP) technique. In this study, we identified an electrophysiological index for the maintenance of tone sequences in ASTM, the Sustained Anterior Negativity (SAN).

Perhaps because it was previously often implicitly assumed that ASTM could be understood via the study of verbal material, few studies have examined the neurophysiological correlates of ASTM, and even fewer have avoided contamination by verbal or phonological codes. Pratt, Michalewski, Barrett, and Starr (1989) had their participants perform a Sternberg task using either visual or auditory digits, or musical notes. They found that only the musical notes showed a negative shift in amplitude correlated with increasing set size at parietal site Pz, while all stimulus types yielded a statistically significant negative shift at frontal sites. However, they analysed their data during and after probe presentation, which makes the attribution of the results to STM processes difficult. Indeed, in the Sternberg task, each trial has three phases: presentation of a memory set, a retention interval, and comparison of a probe to memory. The task was to decide whether the probe was or was not in the memory set. At the onset of the probe, many processes are engaged: perceiving and encoding probe stimulus, memory maintenance, memory scanning and matching to the probe, response selection, and motor preparation. Without specific experimental manipulations designed to isolate individual underlying mechanisms, it is nearly impossible to determine what aspects of the activations recorded during this time period are a reflection of STM processes, and what aspects are due to other task-related, but not specifically memory-related processes. In a follow-up study, Pratt, Michalewski, Patterson, and Starr (1989) focused their analyses on activations occurring during and after the memory set presentation of a Sternberg task.
in which they varied memory load (i.e., the number of to-be-remembered items). They observed a positive parietal wave following the presentation of tones but not following digits presented acoustically, but the same wave was observed for digits presented visually. However, in their analysis they averaged together the EEG signals of several stimuli presented sequentially: for example, in the Load 5 condition, they averaged together the first, third, and fifth stimuli presented in the memory set. This means that the period that followed stimulus presentation sometimes contained further stimulation (the second or fourth stimulus, in our example), and therefore possibly, processes (like perception) that are not part of STM (see Kaufman, Curtis, Wang, & Williamson, 1992, for a similar issue with MEG data). Zatorre, Evans, and Meyer (1994), using PET, found activations in several regions, including the right temporal and frontal lobes in a pitch discrimination task in which trials with and without interference were contrasted. Given that the whole trials were analysed, the authors hypothesised that the many areas activated during the task (relative to a passive listening task) might reflect the many processes required by the task, including retention in STM. Gaab, Gaser, Zaehele, Jancke, and Schlaug (2003) based on fMRI suggested that the supramarginal gyrus and the dorsolateral cerebellum might be involved in ASTM. However, they contrasted the activations obtained in their memory task with a silent condition. Therefore, it is unclear whether the observed activations are the result of sensory stimulation, encoding, retention in ASTM, or some combination of these different processes.

One study, that of Lang, Starr, Lang, Lindinger, and Deecke (1992), has avoided the potential pitfalls of concurrent processes by focusing on the retention interval of a delayed match-to-sample task. In this task, participants were presented with a sequence of items to memorise, which were followed, after a retention interval, by a second sequence. The task was to indicate that if the two sequences were the same or not. This procedure ensured that retention results would not be contaminated by stimulation, response-selection, or other processes concurrent with STM processes. The authors studied the slow EEG potentials between the memory and probe sets of a Sternberg-type task involving auditory digits, visual digits, and tones. Their goal, much like ours, was to distinguish the neurophysiological correlates of the maintenance in STM of stimuli presented in different modalities in STM. The most important difference they found was between the visual and auditory presentation of digits. However, they also found increased negativity at frontal electrode sites F3 and F4 for tones compared to auditorily presented digits at the beginning of the retention interval. This suggests a possible special role of brain activity projecting to frontal electrodes for non-verbal auditory retention interval. This suggests a possible special role of brain activity projecting to frontal electrodes for non-verbal auditory retention is reached or exceeded. In other words, if an individual cannot retain any more stimuli because memory is full, then maintenance-related activation should not increase further even when more stimuli are presented (Robitaille et al., 2010; Todd & Marois, 2004; Vogel & Machizawa, 2004). That is, the brain activity we wish to measure should reflect the functional load in memory and not the stimulus or sensory load. Moreover, this should be true not only at the group level but also at the individual-participant level: the increase in activation with load should plateau earlier for participants with a lower STM capacity than for those with a higher STM capacity. Following this third principle, we correlated a measure of individual STM capacity, K (see Cowan, 2001; Pashler, 1988), with individual variations in electric brain activity with memory load.

A final aspect of our approach to isolate ASTM from other brain processes was to use simple sound stimuli that engage as few memory systems as possible. Although the use of verbal stimuli is necessary if one wants to study verbal memory (see D’Esposito, 2007, for a review), the use of such stimuli in the study of basic mechanisms of acoustic memory is problematic. Such stimuli will likely engage multiple underlying processing and memory systems...
rendering the isolation of basic acoustic memory mechanisms essentially impossible. We used simple pure tones with frequencies chosen not to belong to standard musical scales. We will refer to them as ‘non-musical’ tones (Trehub, Schellenberg, & Kamenetsky, 1999). Contrary to spoken or sung words or syllables as well as familiar sounds, non-musical tones cannot be readily labelled or recoded into other forms of representation, and thus should be mainly retained in a low-level, acoustic format.

These principles were first applied to an ASTM task. Two, four, or six non-musical tones (preceded by white noise fillers to equate total stimulation duration between load conditions) were presented binaurally, as illustrated in Fig. 1. After a silent, 2-s retention interval, again two, four, or six tones were presented (this time, without the white noise fillers), and participants had to decide if the second sequence was the same as the first or not. Results from the memory task were compared with those from a control task in which exactly the same initial stimuli were presented, but with instructions to ignore the first sequence and instead decide if the last two tones of the second sequence (generated independently from the first) created a rising or falling pitch change. In the control task, no stimuli had to be retained during the interval between the first and second sequences. Importantly, there was no incentive to encode and remember the first sequence during the 2-s interval because it provided no useful information for the task to be performed with the second sequence, which was constructed independently from the first. Thus, the memory and control tasks were equated on all low-level stimulus presentation parameters up to the presentation of the second sequence, allowing a direct comparison between them during the retention interval. Hence, any brain activity specifically related to the retention of acoustic properties of sound should be present during the memory task but absent during the control task, particularly in the latter portion of the 2-s silent period between the first and second sequences of tones.

Our hypothesis was that activity related to the maintenance of representations in ASTM should increase as memory load increased. Furthermore, the degree of increase should predict individual differences in memory capacity across individuals. We assessed this relation by correlating changes in brain activity across memory load with a measure of ASTM capacity derived from accuracy in the memory task.

2. Experiment 1

2.1. Method

2.1.1. Participants

Thirty-nine participants took part in the experiment. Data from one participant was rejected because it was found after testing that this person suffered from tinnitus. Mean age of the participants was 24.6 years (range: 20–33). Nineteen of them were women and all but one were right handed. All remaining participants reported normal hearing and no neurological problems.

2.1.2. Materials

The stimuli consisted of one white noise stimulus and 14 non-musical tones. By non-musical, we mean tones that had a pitch that did not correspond to the equal-tempered scale. The 14 tones ranged from 380 to 1375 Hz. Their exact frequency (380, 419, 463, 511, 564, 623, 688, 759, 838, 925, 1022, 1128, 1245, and 1375 Hz) was determined by multiplying the starting frequency and each resulting frequency by $2^{1/7}$ (1.104), so that the differences between the tones were equal on a log scale with each octave divided into seven equal steps. All sounds were represented digitally with 16 bits of resolution at a sampling rate of 44.1 kHz and had a duration of 100 ms (including 10-ms rise and fall times), and were presented binaurally at 65 dB SPL via etymotics ER-1 earphone and soft plastic ear inserts.

2.1.3. Design and procedure

A 2 × 3 within-subject design was employed with two repeated-measures factors: Task (memory vs. control) and Load (2, 4, and 6). Whereas the two tasks were blocked, load varied randomly from trial to trial. Participants performed a single session consisting of 600 trials, separated into 10 blocks of 60 trials. Participants performed five blocks in each task (300 trials per task), alternating between the memory task and the control task. The task performed in the first block was counterbalanced across participants. Prior to the experimental blocks were two blocks of practice trials, one for each task.

2.1.3.1. Memory task. Participants initiated each trial by pressing the space bar. After an average delay of 500 ms (± 100 ms random jitter across trials), a fixation cross appeared at the centre of the screen, and participants were instructed to maintain fixation on it and avoid blinking for the remainder of the trial. The fixation cross remained on the screen until 500 ms after the end of the second sequence. A first sequence of seven sounds was presented 500 ms after the onset of the fixation cross. The last 2, 4, or 6 sounds were non-musical tones varying in pitch, preceded by 5, 3, or 1 white noise fillers, respectively. The tones were selected randomly on each trial and were never repeated in a sequence. Each stimulus was separated from the next by a 100-ms silent interval. Participants were told to ignore the fillers and memorise the tones. The first sound sequence was followed by a 2000-ms silent retention interval. After the retention interval, a second sequence composed of the same number of pure tones as in the first sequence was presented, this time with no white noise fillers; this was done to ensure participants would not be encouraged to encode fillers. Example trials are illustrated in Fig. 1. The test

Fig. 1. Examples of sequences in the memory task of Experiment 1. Stimulus frequency is represented by the height of the columns; white noise sounds are represented by pixelized columns. In the Load 2 condition (a), the two tones in the memory set are preceded by five white noise bursts. In the Load 4 condition (b), the four tones are preceded by three white noise bursts, while in Load 6 (c), the six tones are preceded by a single white noise burst. No white noise bursts are presented in the probe set.
Independent Component Analysis (ICA) was performed on the data of participants. Apparent artefacts not removed by the preceding procedure were removed manually if they were observed over a period of 50 ms, or if it varied by more than 200 μV over a 150-ms period. For half the participants, a ‘same’ response was indicated by pressing the ‘C’ key, whereas for the other half of the participants, this was reversed. Feedback was provided 500 ms after the response. The circle turned green if the response was correct, bright red if it was incorrect, and dark red if the participant did not respond within 3000 ms. Feedback remained on the screen until the participant initiated the next trial.

2.1.3.2. Control task. The general structure of stimulation sequences in the control task was as follows: however, participants were told to ignore the first sequence and to indicate whether the last tone of the second sequence was higher or lower in pitch than the second-to-last tone. Contrary to the experimental task, the two stimulus sequences were independently created random sequences, and thus the test sequences were not related to the initial sequences. Hence, the first sequence provided no indications about the correct response and had no task relevance. Half the participants responded by pressing the ‘C’ to indicate a ‘lower’ response (falling pitch across the last two tones) and ‘M’ for a ‘higher’ response (rising pitch). The opposite mapping was used by the other half of the participants.

When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order. When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order. When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order. When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order. When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order. When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order. When a participant failed to reach 70% accuracy in the control task, he or she was removed from analyses. The reason for this exclusion criterion was that the memory task requires encoding of auditory stimuli in a particular order.

2.1.4. EEG recording and analyses

EEG signals were recorded using a Biosemi ActiveTwo system, at a sampling rate of 512 Hz, but was downsampled to 256 Hz to speed up analyses. The location of each active Ag/AgCl electrode on the 64-electrode cap followed the 10–20 system. External electrodes were placed on the left and right mastoids. The horizontal electrooculogram (HEOG), defined as the voltage difference between the two canthi electrodes, below the left eye, and on the left and right mastoids. The horizontal electrooculogram (HEOG), defined as the voltage difference between the two electrodes placed above (Fp1) and below the left eye, was used to measure horizontal eye movements, while the vertical EOG (VEOG), defined as the voltage difference between the electrodes placed above (Fp1) and below the left eye, was used to detect eye blinks. The signals were filtered offline to the average of the two mantids.

A 0.05-Hz, 12–dB/octave—20–Hz, 24-dB/octave band-pass filter was applied offline to signals from all electrodes to remove slow drifts and high frequency noise. A 5-Hz, 48-dB/octave, low-pass filter was also applied to the HEOG and VEOG channels only, to smooth them in order to facilitate the EEG artefact rejection process. Signals from all channels were removed on all electrodes when the HEOG signal varied by more than 35 μV over a 200-ms period, or if the VEOG signal fluctuated by more than 50 μV over a 150-ms period.

Data from any channel were removed from a trial if a 100-μV variation was observed over a period of 50 ms, or if it varied by more than 200 μV in a 500-ms period. Data were segmented by trial in periods of 3600 ms. Each segment started 200 ms before stimulation began, and finished 100 ms after the onset of the second sound sequence. Only the first sound sequence was therefore included in the segments. The data were baseline corrected from 200 to 0 ms before stimulation began. Artefact-free data were averaged separately for each task and load condition.

In order to visualise the overall scalp distribution of the effect of load, we subtracted the mean voltage in the Load 2 condition from the mean in the Load 6 condition, at each electrode, and produced a voltage map for each task. These maps are shown in Fig. 2. The strongest load-related response in the memory task was concentrated at fronto-central electrodes in the region of AFz, with smaller and non-systematic variations in the control condition. We focused our analyses on fronto-central recording site AFz and the eight electrodes surrounding it, where variations of Load were near the maximum of the component visible on the scalp distribution (see Fig. 2).

Grand average waveforms for each Load and Task combination are shown in Fig. 3, for electrode AFz, Fp1, Fpz, Fp2, AF3, AF4, F1, F2, and Fz. The initial portion of the waveforms show the sensory responses produced by the presentation of the first white noise burst, in all conditions, followed by either more white noise bursts that produced smaller subsequent responses or pure tones that produced larger responses. For example, the waveform for the Load 6 conditions shows clearly the six tone-elicited auditory N1 responses expected at fronto-central electrodes.

As can be seen in Figs. 2 and 3, increasing memory load caused a greater negativity at fronto-central electrodes in the memory task, and a weaker (and if anything opposite) response in the control task. We quantified the memory—load response by computing the mean voltage in each condition between 900 and 1900 ms. We focused on this latter portion of the retention interval to ensure that we did not include any late sensory activity caused by the mere presentation of the tones, and we excluded the last 100 ms of the retention interval to reduce possible effects
related to anticipation of the second sequence. The mean amplitude of AFz, which was nearest to the peak of the scalp distribution for the component of interest (as can be seen in Figs. 2 and 3), was submitted to a repeated-measures ANOVA with factors Load (2, 4, 6) and Task (memory, control). We found a significant Load × Task interaction, \( F(2, 44) = 3.37, \text{MSE} = 7.98, p = 0.043 \), that reflected a significant effect of Load in the memory condition, \( F(2, 44) = 4.90, \text{MSE} = 6.30, p = 0.012 \), but not in the control condition, \( F(2, 44) = 0.32, \text{MSE} = 10.38, p = 0.71 \). The main effect of task, namely a greater overall negativity in the memory task than in the control task, was reliable, \( F(1, 22) = 4.62, \text{MSE} = 12.27, p = 0.043 \), but there was no main effect of Load, \( F(2, 44) = 0.84, \text{MSE} = 8.71, p = 0.42 \). In short, larger loads increased negativity at fronto-central site AFz, but only in the memory task. In order to verify that these patterns were not confined to a single electrode, we also performed a similar analysis on the pooled response for AFz along with eight neighbouring electrodes (Fp1, Fpz, Fp2, AF3, AF4, F1, Fz, and F2). The ANOVA confirmed what was clearly visible from observation of the waveforms in Figs. 2 and 3, namely that the patterns seen at AFz were also present nearby, but with slightly lower amplitudes. This was reflected in the ANOVA by an almost significant Load by Task interaction, \( F(2, 44) = 2.94, \text{MSE} = 5.73, p = 0.064 \). Task also showed a significant effect, \( F(1, 22) = 4.72, \text{MSE} = 8.62, p = 0.041 \), with a more negative mean overall in the memory task than in the control task. No other effect was significant, all \( Fs < 0.78 \), all \( ps > 0.45 \). In short, larger loads increased negativity at fronto-central site AFz, but only in the memory task.

To examine the relationship between the sustained anterior negativity found in the memory task and individual memory capacity, we correlated the change in voltage as a function of load and a behavioural measure of ASTM capacity. We estimated memory capacity for each participant by computing the memory capacity index \( K \) (see Cowan, 2001, adapted from Pashler, 1988), in which \( K = S(H-F) \), where \( S \) is the number of stimuli, \( H \) is the hit rate, and \( F \) is the false alarm rate. The mean \( K \) as a function of Load is presented in Fig. 4. Note that memory capacity increased significantly between Load 2 (mean = 1.9 items) and Load 6 (mean = 3.75 items), as confirmed by a repeated-measure one-
way ANOVA, $F(2, 44) = 60.80, p < 0.0001$. A Tukey-A test also confirmed that memory capacity increased significantly from Load 4 (mean = 3.3 items) to Load 6, $p < 0.05$. These results, as well as the amplitude waveforms shown earlier, suggest that if we have reached our subjects’ maximum capacity, we have not surpassed it, as in both cases we have not reached an asymptote.

Individual maximum $K$ values were correlated to the difference in mean amplitude across Loads 6 and 2. An individual with a lower capacity should have a smaller $K$ and a smaller voltage change across Loads 6 and 2, whereas an individual with a larger $K$ was expected to show a larger increase in fronto-central negativity from Load 2 to Load 6. Scatter plots showing the relationship between voltage change (across loads) and capacity are shown in Fig. 5. Because we expected the results to go in one particular direction (negativity increasing with larger memory capacity), we used one-tailed $p$ values. As memory capacity ($K$) increased, the voltage difference across Loads 6 and 2 became more negative, producing a significant negative correlation, $r = -0.40$, $t(21) = -2.00$, $p = 0.029$, for electrode AFz, as well as a significant negative correlation, $r = -0.52$, $t(21) = -2.80$, $p = 0.009$, for the pooling of electrodes Fp1, Fpz, Fp2, AF3, AFz, AF4, F1, Fz, and F2. These findings provide strong converging evidence for our interpretation of the observed sustained anterior negative ERP as a reflection of maintenance in ASTM.

2.3. Discussion

The goal of Experiment 1 was to identify neural activity specifically related to the retention of pitch information in ASTM. We developed a new paradigm based on the presentation of sequences of sounds equated for overall length (here all sequences had seven events) but in which we varied the number of critical items to be encoded into ASTM (2, 4, or 6 items). Participants remembered these sequences of non-musical tones for a later memory test, and we focused on brain activity recorded during the intervening silent retention interval.

Following initial sensory-encoding responses, we observed a sustained negative response at fronto-central electrodes that persisted through the entire duration of the retention interval. We called this ERP the sustained anterior negativity (SAN), to reflect the approximate topography, time-course, and polarity of the signal. The SAN increased in amplitude with memory load (Principle 2), providing one important piece of support for the notion that our approach enabled us to isolate brain activity that is specifically related to maintenance in ASTM. A second important piece of support is that this load-dependant activity was measured during the retention interval (Principle 1), at which time there was no stimulation, and all encoding processes had finished, and prior to retrieval, comparison, and response-selection operations had begun. The stimuli also enabled us to isolate ASTM, per se, by limiting possibilities of activation in other memory systems that were not targeted by our research (musical/melodic, phonological, lexical, semantic, verbal, graphemic, etc.; Principle 4).

Another important result was that changes in the amplitude of the SAN as memory load increased were significantly correlated with individual-participant’s ASTM capacity estimated from behavioural performance in the memory task. Participants with a smaller capacity would not be expected to be able to encode as much additional information into ASTM as those with a larger capacity, as Load increased from 2 to 6. This should translate into a proportional change in SAN amplitude. For example, a participant with a capacity of four items should show an increase in SAN when Load changed from 2 to 4, but should have no further increase in SAN when presented with six items. Such a participant would thus be expected to have a smaller change in SAN from 2 to 6 compared with another participant who had a capacity of 6 and for whom the SAN would continue to increase beyond four items. Thus, we predicted that changes in the SAN across Load conditions should be larger for participants with a larger capacity. This is exactly what we found, suggesting that the SAN is not only capable of indexing differences in the degree of involvement of ASTM at the group level, but also of predicting individual differences in ASTM capacity. This, in turn, provides strong evidence for a functional interpretation of the SAN as an index of brain activity reflecting maintenance in ASTM.

Importantly, no load-dependent activity was measured during the silent interval in the control task. These results enable us to
rule out a number of alternative accounts of the load-dependent brain responses observed during the memory task. In particular, we can rule out any account that hinges on the notion of persistence of sensory activation and differences of persistence across load conditions, or of automatic encoding and retention of stimuli. The absence of load-dependent variations in SAN in the control task suggests that maintenance in ASTM involves an active process without which representations in ASTM are quickly lost (or perhaps not encoded in the first place).

Our results share some similarity with those gathered by Pratt, Michalewski, Barrett, et al. (1989) and Pratt, Michalewski, Patterson, et al. (1989) in that both studies found a negative-going shift increasing with set size. However, in the studies by Pratt and colleagues, the negative shift linked exclusively to sound stimuli was found at posterior site Pz, whereas we found no such activity at site Pz or near it. This discrepancy could result from the inclusion, by Pratt and colleagues, of stimulation as well as the retention interval in the time interval included in their ERP analyses.

Close inspection of the early (stimulation) portion of waveforms in Fig. 3 shows interesting fluctuations in amplitude. As mentioned earlier, we can clearly see an N1 and a P2 associated with each stimulus presentation (with smaller N1s for white noise). We can also see a general negative-going trend that increases with each new sound. At the end of stimulation, the amplitude at Load 6 reaches about ~6 μV, and at Load 4 it reaches about ~5 μV. This negativity is interesting as it seems incremental, increasing after each tone presentation. It may be related to encoding and the beginning of the maintenance of an increasing number of stimuli, but our results cannot provide definitive evidence in this regard. We leave in depth analysis of this period to further studies. Nonetheless, at the end of the stimulation period, amplitudes are already negative. One could suggest that our results are simply this negativity that is carrying on. We are of the opinion that stimulation alone cannot explain our results for various reasons. First, we can clearly see an N1 and a P2 associated with each stimulus before. We will therefore measure the SPCN during the retention interval of our sequential VSTM task, as well as examine activations at fronto-central areas to verify if we observe a SAN in memory task without which representations in ASTM are quickly lost.

3. Experiment 2

An interesting aspect of the SAN observed in the memory task of Experiment 1 was that the SAN scalp distribution was different from the one observed during the retention of low-level visual features (such as colour or line orientation). The sustained posterior contralateral negativity (SPCN) has been observed in visual experiments that conform to the research principles we adopted for the present research, and consistently has a very posterior distribution (e.g., Brisson & Jolicœur, 2008; Jolicœur, Brisson, & Robitaille, 2008; McCollough, Machizawa, & Vogel, 2007; Prime & Jolicœur, 2010). These very large differences in scalp distribution suggest that the underlying neural generators are rather different and thus that the two memory systems (ASTM and VSTM) have at least some generators that are unique to each memory system.

Recent studies investigating VSTM have been mostly interested in lateralized potentials like the SPCN, and they often report analyses only for this kind of data. As a result of this focus on the SPCN, it is possible that a non-lateralized fronto-central negativity associated with load increase was overlooked. We present, in Fig. 6, scalp topography results from a VSTM experiment we published previously (Perron et al., 2009). The topographic map shows mean voltages from the lowest memory load (two coloured bars presented laterally) subtracted from the voltages observed at the highest memory load (three coloured bars also presented laterally) without regard to the visual field of the stimuli to be encoded. Although Perron et al. (2009) found a clear lateralized potential (SPCN) that increased with set size, reported in the original paper, we could see no increase in fronto-central negativity with set size in the non-lateralized waveforms. These results suggest that the SAN observed in Experiment 1 of the present study likely reflects neural generators related to ASTM rather than to other, more general, aspects of the task, such as general task difficulty. General difficulty increased in the Perron et al. study as memory load increased, and yet no SAN was found in that study. The re-analysis of the Perron et al. study thus suggests that VSTM and ASTM are mediated by distinct generators (Fig. 6).

There is, however, an important procedural difference between our ASTM task and typical VSTM tasks, such as the one used by Perron et al. (2009), that could explain the difference in the results obtained. Most VSTM tasks present stimuli simultaneously, whereas stimuli were presented sequentially in our ASTM task. It is therefore possible that the component we have identified is an index of the active maintenance of stimulus order or of sequences of stimuli, regardless of their nature, or that adding a temporal dimension in the stimuli to be remembered engaged entirely different sets of processes. If that was the case, then a similar task, using sequential stimuli in a different modality, would yield similar results.

To test for this possibility (and others discussed in subsequent sections), we created a sequential VSTM task equivalent to Experiment 1, using coloured disks as stimuli instead of tones. As in other VSTM tasks, the target stimuli were presented laterally, because we wanted to measure the SPCN component. To our knowledge, the SPCN has not been observed with sequential stimuli before. We will therefore measure the SPCN during the retention interval of our sequential VSTM task, as well as examine activations at fronto-central areas to verify if we observe a SAN in the visual modality. This task being analogous to Experiment 1, it should provide satisfactory answers to many pending questions: if the component we found is the reflection of general task difficulty, then change in stimulus modality should not affect the waveforms measured. Finally, if we have found an index of the maintenance of a sequence of stimuli, then we should also observe results similar to those in Experiment 1, as the visual stimuli in this experiment were presented sequentially.

3.1. Method

3.1.1. Participants

Thirty-five students from Université de Montréal took part: 21 were women and 25 were right handed. Their mean age was 21.53 years, the youngest participant being 19, and the oldest 28. All had normal or corrected-to-normal vision.

3.1.2. Stimuli

The stimuli used in the experiment consisted of coloured disks that subtended a viewing angle of 2.1° from the participants’ seating position, which was 57 cm from the screen. They could be one of the six different colours, namely salmon pink, light green, light blue, olive green, lilac, or light grey, and were all approximately...
Applied to the data of all participants except three, for which we used a highpass filter of 0.10 Hz instead of 0.05 Hz, to attenuate large-amplitude slow waves contaminating the data.

EEG was recorded at a sampling rate of 512 Hz and later downsampled offline at 256 Hz to speed up analysis. Segments with a VEOG greater than 50 μV were segmented as well as segments with HEOG activity change of more than 35 μV in a 200 ms window. Then, if activity at electrodes AF3, AFz, AF4, F1, F2, Fz, FP1, FPz, P2, (critical electrodes in Experiment 1) or PO7, PO8, P7, P8, P30, or P40 (critical electrodes when measuring the SPCN) varied by more than 100 μV over a period of 50 ms, or if it varied by more than 200 μV in a 500-ms period, data from all channels were removed for that period. If data from another channel exceeded the above thresholds, data from that channel only were removed for that period. Apparent artefacts not removed by the preceding procedure were removed manually after visual inspection. The data from one participant was lost due to technical reason, while another participant did not have enough time to complete the whole experiment. The data from 10 other participants were rejected from analyses because they systematically moved their eyes towards the target, ending up with very few valid trials. This left 23 participants included in the following analyses. All these participants had a residual HEOG of less than 7 μV when subtracting the mean HEOG from left-side targets from the mean HEOG of right-side targets. Twenty-one out of 23 participants had a residual HEOG subtraction of less than 3 μV. The average residual effect, averaged over the period ranging from −1700 to 1900 ms, was −113 μV, corresponding to a mean eye movement towards the targets of less than 0.07° of visual angle.

3.2. Results

3.2.1. Behavioural data

To establish how Experiment 2 compared to Experiment 1, we first take a look at the accuracy in Experiment 2. The task proved to be difficult in general, as mean accuracy dropped from 87% at Load 2 to 72% at Load 4. It was near perfect (99%) at Load 0, as
there was no specific visual memory task in this case. A matched-pairs t-test was performed on accuracy at Loads 2 and 4. The t-test confirmed that accuracy at Load 2 was larger than accuracy at Load 4, *t*(22) = 9.36, *p* < 0.0001.

3.2.1.1. Electrophysiological data. The data were segmented into 4200-ms epochs, starting 2100 ms prior to the retention interval (the beginning of the retention interval was time zero), and finishing 100 ms after the onset of the second sequence, 2100 ms after the start of the retention interval. Each trial was baseline-corrected using the 200-ms interval before stimulation started (i.e., −2100 ms to −1900 ms relative to the beginning of the retention interval). On the right of Fig. 8 we display the waveform for each load at AFz. Activity at AFz was quite different from that observed in Experiment 1: here we observed positive deflections, not negative. Importantly, this activity did not change with memory load, unlike what was found in Experiment 1. The topographic map on the left side of Fig. 8, which shows the subtraction of Load 2 mean voltages from Load 4 mean voltages, revealed no portion of the map suggesting a load-related voltage change. These results confirm those of the re-analysis of the Perron et al. (2009) data shown in Fig. 6. To confirm our observations, we ran an ANOVA on the nine fronto-central electrodes for the factor Load, (contralateral minus ipsilateral), averaged in a window from 900 to 1900 ms. As expected from what can be seen in Fig. 8, there was no effect of Load, *F*(2, 44) = 0.51, *MSE* = 3.74, *p* = 0.620. A similar ANOVA was run on the single electrode AFz, which showed peak amplitudes in Experiment 1. Without surprise, this ANOVA does not show any sign of Load effect, *F*(2, 44) = 0.58, *p* = 0.563.

But can the maintenance of a sequence of visual stimuli yield an SPCN during the retention interval of a STM task as it does for simultaneously presented visual stimuli? To find out, we calculated the SPCN waveforms at each load by averaging voltages measured at sites ipsilateral to the encoded stimuli (recall that target sequences were presented bilaterally but that participants were instructed to encode stimuli only from one side or the other) and then subtracting them from voltages averaged at sites contralateral to the encoded stimuli. We show a topographic map of the subtraction of the SPCN values at Load 4–Load 2 for the end of the retention interval, namely 900–1900 ms, in the left panel of Fig. 9. Because the SPCN calculation gives an image of half a scalp, the measurements were mirrored to create a complete, symmetric map. Also, because the SPCN is observed at posterior sites, we focused on a map of the back of the head.

This map revealed a greater contralateral negativity at electrodes P07/P08, indicating that the SPCN increased with load. In the right panel of Fig. 9, we show the SPCN waveforms at electrodes P07/P08, for Loads 0, 2, and 4. The SPCN started about 400 ms after the onset of stimulation, which is not surprising, as participants most probably started to maintain the items as soon as they were encoded. This is true for Load 4 only, as Load 2 does yield a very small SPCN. This is similar to what was observed for the SAN in Experiment 1 in which the retention of two tones produced only a very small SAN deflection. The SPCN remained stable for the duration of the retention interval. We submitted SPCN values (contralateral minus ipsilateral), averaged in a window from 900 to 1900 ms after the start of the retention interval, for each memory Load, at sites P07/P08, to a repeated-measures ANOVA with Load (0, 2, or 4 items) as the single within-subjects factor. This ANOVA revealed a significant effect of Load, *F*(2, 44) = 3.33, *MSE* = 4.03, *p* = 0.04, indicating that the SPCN increased with memory load. To verify further that this increase in amplitude was linked to the number of items memorised, we checked if there was a correlation between individual participants’ memory capacity (K) and the difference in SPCN amplitude between Loads 4 and 2. This correlation was significant, *r* = −0.46, *t*(21) = −2.37, *p* = 0.014, one-tailed, showing that participants who could retain more items also had a larger increase in SPCN amplitude with load than participants who could retain fewer items (see Fig. 10).

3.3. Discussion

We found no increase in negativity of the electrophysiological signal at fronto-central sites as memory load was increased, or at any other site, in Experiment 2, meaning that we did not observe a SAN in Experiment 2. However, we observed evidence of increased activity at posterior recording sites linked to larger VSTM loads in the form of an SPCN component, which is in conformity with the existing literature on VSTM. Moreover, sequential stimulus presentation did produce activations similar to those found in the usual, simultaneous presentation paradigm used in VSTM tasks, by showing increases in SPCN with load, and also a significant correlation between individual participants’ performance and the increase in SPCN amplitude measured.

More generally, the results support the hypothesis that Experiments 1 and 2 engaged different memory mechanisms during the retention interval and hence that the retention of auditory and visual representations is mediated by distinct neural mechanisms.

4. General discussion

In this paper we provided evidence for neurophysiological correlates specific to ASTM for sequences, and dissociated them from those of VSTM. In Experiment 1, we showed that electrophysiological signals at fronto-sites increased with ASTM Load, and that this increase predicted individual participants’ memory capacity. In Experiment 2, we showed that maintenance of sequences of visual items did not yield a SAN, but produced a load-dependent SPCN. Similarly to previous findings with simultaneous visual presentations (Vogel & Machizawa, 2004), the increase in SPCN amplitude with increasing memory loads was correlated with the number of visual items participants can recall, providing strong evidence that the SPCN is related to the maintenance of items in VSTM.

The findings of this paper provide evidence in favour of a decentralised view of STM for sensory stimuli (D’Esposito, 2007; Postle, 2006). It provides indirect support for the hypothesis that sensory memory is maintained in the structures responsible for the perception of the same stimuli, as the fronto-central recording sites we found are compatible with frontal and temporal activations. It also provides direct support for this thesis by showing that visual and auditory stimuli are most likely maintained in different structures, as maintenance activity is observed at very different electrodes for the two modalities. EEG activity recorded at fronto-central sites, like the one observed in Experiment 1, and that of Guimond et al. (2011), are generally compatible with the ASTM-related activations measured in a number of areas in the brain via MEG and fMRI. The SAN could be related to the activations found by Grimault et al. (submitted for publication) in an ASTM study that used a very similar paradigm as the one in Experiment 1, applying the same principles that we used in the present study. Using MEG, Grimault et al. identified several brain areas that showed activation correlated with individual memory capacity in the frontal and temporal cortex, namely the left superior/middle temporal gyrus, the left pre/postcentral gyrus, the left precentral/middle frontal gyrus, the right inferior frontal gyrus and the right middle frontal gyrus. These results only partially replicated other ASTM imaging results, likely because other studies did not implement all the design principles we followed, and therefore perhaps measured functional images of more than maintenance in ASTM (see for example Crottaz-Herbette et al., 2004; Gaab et al., 2003; Klostermann, Loui, & Shimamura, 2009; Koelsch et al., 2009; Strand et al., 2008).
The combination of results across Experiments 1 and 2 allows us to rule out many alternative hypotheses concerning the neural activity reflected in the SAN. First, we verified that the SAN was not the reflection of the processing of a sequence of stimuli, irrespective of stimulus modality. Recall that in most VSTM studies, stimuli are presented simultaneously, whereas in our ASTM task (Experiment 1), stimuli were presented sequentially. Therefore, the SAN could reflect this difference in paradigm rather than a difference in stimulus modality. The absence of a SAN in the sequential VSTM task of Experiment 2, however, combined with an increase in SPCN with memory load in the same experiment, is evidence that this was not the case. Further evidence is provided by other work from our lab. In the Guimond et al. (2011) study, simultaneous presentation of one or two tones was used in an ASTM task that also produced a load-sensitive SAN, suggesting further that maintenance or processing of a sequence is not at the root of this effect.

The results also allow us to reject the possibility that the increase in activation recorded at fronto-central sites in Experiment 1 was due to an increase in general task difficulty, rather than to increased activity in mechanisms specific to the maintenance of acoustic information in ASTM. For instance, Ruchkin, Johnson, Mahaffey, and Sutton (1988) found an increased negativity at posterior and central sites when they increased task difficulty linked to the visual presentation of digits. They attributed the slow waves they observed to the longer processing period associated with the more difficult tasks: it should be more difficult, and therefore take more time to divide numbers than to subtract them, and take more time to subtract two numbers than to maintain one. This was reflected in their results. However, we have little evidence this is the case in our task. Retaining more items might be more difficult, but the sustained nature of maintenance, and its completion, is decided by external factors (in our case the length of retention interval) much more than by task difficulty. The sustained nature of our waveforms as well as the topography of the SAN are different from that of Ruchkin et al., making this explanation of general difficulty less attractive. Moreover, we have direct evidence against this explanation in Experiment 2. Despite behavioural data showing that the task was of similar, if not greater, difficulty (74% and 87% at Loads 4 and 2 in Experiment 2, vs. 88% and 95% at Loads 4 and 2, and 78% at Load 6 in Experiment 1), we did not observe a SAN in Experiment 2.

Fig. 8. Results from Experiment 2: left panel. Topographic map of the difference in activation between Loads 4 and 2, in microvolts (μV), averaged during the period from 900 to 1900 ms from the beginning of the retention interval of Experiment 2. Right panel: waveforms at AFz, in microvolts (μV), shown separately for every Load (0, 2, 4), from 2200 ms before retention to 100 ms after the retention interval. Time zero refers to the onset of the retention interval.

Fig. 9. Results from Experiment 2: left panel. Topographic map of the difference in activation of the SPCN between Loads 4 and 2, in microvolts (μV), averaged during the period from 900 to 1900 ms from the beginning of the retention interval. Note that the map is perfectly symmetrical because the SPCN calculation gives an image of half a scalp, and therefore the measurements were mirrored to create a complete map. Right panel. Post-SPCN calculation waveforms at electrodes PO7/PO8, in microvolts (μV), shown separately for every Load (0, 2, 4), from 2200 ms before retention to 100 ms after the retention interval. Time zero refers to the onset of the retention interval.
Clearly, an increase in task difficulty is not sufficient to produce a SAN. Finally, if task difficulty is assumed to increase as accuracy decreases, then our results show a reverse effect. In Experiment 1, participants with higher memory capacity showed larger increases in SAN amplitude whereas participants with lower accuracy showed smaller increases in SAN amplitude, suggesting that the effect observed was not the result simply of more difficult processing.

Experiment 2 helped to investigate whether the SAN was the resultant of activity of a general central executive that would be more solicited whenever a larger memory load is maintained, in any modality (cf. Baddeley, 2000; Baddeley & Hitch, 1974). Because of its 'central' position in the processing chain, this module should be active in a similar manner whatever the type of stimuli used, and therefore, if the SAN reflected the activity of a central executive, it would have been measured in Experiment 2, where the basic task was the same, but the stimuli were different. The results show clearly this was not the case: there was no SAN in Experiment 2, disconfirming the hypothesis that the SAN reflects the activity of a central executive involved whenever task short-term memory load is increased.

Finally, one might wonder if the SAN is a CNV in disguise. However, the SAN is different from the CNV in many ways. First, the CNV is usually defined as a negative-going wave, occurring between the presentation of a warning stimulus (S1) and an imperative stimulus (S2) in a reaction time task. This differs in two ways from the SAN: first, the amplitudes measured during the retention interval of Experiment 1 were initially positive-going, and then mostly stable until the end of the retention interval, contrary to the CNV, which is a sharply negative-going deflection just before the onset of the second stimulus. Also, the increasing loads produced more negative amplitudes, which is contrary to Roth et al.’s (1975) as well as Ford, Roth, Mohs, Hopkins, and Kopell’s (1979) results, which showed that increases in STM load decreased CNV negativity. Secondly, in our task, participants were encouraged to respond as accurately as possible, without pressure to respond quickly. The CNV, on the other hand, is reduced or disappears in the absence of speed pressure (Loveless & Sanford, 1974; Gaillard & Perdok, 1980; Gaillard, Perdok, & Varey, 1980).

The CNV has first been linked to the expectation of the appearance of the S2 stimulus (Walter et al., 1964). If the waveform we observed was a CNV, then we should measure similar negativity, overall, in the memory and the control task of Experiment 1, because in both cases participants anticipated the presentation of the second sequence. Hence, both tasks created expectancy for S2. Moreover, expectancy might be higher in the control task because the only relevant stimuli were presented at S2 in the control task. This is not what we observed, however, activations observed in the control task were much less pronounced than in the memory task. In addition, as accuracy results have shown, the control task was slightly more difficult than the memory task (89% correct in the control task versus 91% correct in the memory task). This difference was not significant, $F(1, 22)=1.99, p=0.172$, but note that many participants could not perform to a level of 70% and were rejected from the analyses. Since S2 discrimination difficulty increases CNV amplitude (Nakamura, Kufui, Kadobayashi, & Kato, 1979), then we have further reason to expect an equal or a larger CNV in the control condition compared to the memory condition. As can be observed in Fig. 3, the mean signal at Fz during the retention interval of the control task was positive-going, and therefore not a CNV, and not larger than in the memory condition despite a more difficult upcoming discrimination. In short, if we measured a CNV in our tasks, then we should be more likely to measure a CNV in the control condition than the memory condition, and we did not, presumably because some necessary conditions for a CNV to be observed was not present in our tasks.

On the other hand, some researchers (Bendixen, Grimm, & Schröger, 2006; Loveless & Sanford, 1974; Rohrbaugh & Gaillard, 1983) suggested that the CNV is in fact the summation of many different components occurring at the same time, and that all of them are not dependent on the contingency of the second stimulus. They identified at least two large families of waves: an early wave that is linked to the presentation of the stimuli, sometimes referred to as the O-wave, for orientation processes, and a second, later family of waves relating to response preparation, sometimes referred to as the E-wave. The portion of the wave we measured would have to correspond to the later wave. However, we examined the activity during memory retention prior to the presentation of the memory probe, and hence before any response preparation was possible. This means that if the waveform we measured was a CNV, it could not be linked to response preparation. This would be a third family of underlying waves. Participants could not know which finger would be pressed until the second stimulus sequence was presented. It is unlikely that motor preparation is at the root of SAN activation observed prior to the presentation of the memory probe sequence. Hence, preparation, whether motor or response, cannot be at the root of the SAN activation. Bendixen et al. (2006) also point to the presence of a timing component to the CNV, a negativity that is present at central and anterior electrodes and which is related to the length of the interval between stimulations. Importantly, the interval length was constant in all our conditions, and therefore cannot explain the difference we observed between the control task and the memory task. One might argue that the interval actually starts with the presentation of the first relevant stimulus, and therefore the interval would be longer in the Load 6 condition than in the Load 4 or Load 2 conditions. In their experiment, Bendixen et al. (2006) varied the response stimulus interval. Instead of an increased CNV, they observed a increase in positivity in one of their conditions, an increase in parietal negativity in another, suggesting increasing response to stimuli intervals do not necessarily lead to an increase in fronto-central negativity, and can even lead to the reverse. In the one (out of four experiments) in which they found an increasing negativity with increasing interval duration, all conditions reached a maximum value at less than 1000 ms. Because our retention interval was twice as long as this value, it is unlikely that the systematic differences we found across memory load conditions reflect the same process producing their results (which were in a very different task).

Finally, we reiterate that expectancy and preparation were very similar in Experiments 1 and 2. If present in Experiment 1, a CNV

![Fig. 10. Scatter plots relating individual memory capacity ($K$) and difference in the mean amplitude observed within the 900–1900-ms period of the retention interval between Loads 4 and 2 for the SPCN values calculated at electrodes PO7 and PO8.](image-url)
should also be present in Experiment 2, as the contingencies were the same. However, there was no negative frontal wave measured in Experiment 2, which provides strong evidence that the memory-related SAN wave measured in Experiment 1 was not a CNV.

5. Conclusion

In summary, our principled approach to the study of ASTM allowed us to isolate a new ERP component, the SAN, which indexes the maintenance of representations of acoustic information in ASTM. The SAN increased in amplitude with larger memory loads during the retention interval of the acoustic memory task, but not in the control task. The SAN also predicted individual differences in memory capacity, providing additional strong converging evidence for our functional interpretation of the component. Following the same principles, we also investigated the retention of a sequence of visual items in STM. We reproduced results obtained in previous studies using simultaneous presentation of the visual items in the form of the SPCN component that increased with load, and predicted individual differences. Most importantly, we found no SAN in the VSTM task, thus providing strong evidence that the SAN is a specific index of maintenance in ASTM. We anticipate that this component will be useful in further studies of the role of ASTM in cognition, in understanding brain plasticity during development and as a result of learning (e.g., musical training), and in tracking consequences of trauma and pathology for the retention of acoustic information.

Acknowledgements

This work was supported by grants from the Canadian Institutes for Health Research, the Canada Foundation for Innovation, the Canada Research Chairs Program, and by the Fonds de recherche du Québec - Nature et technologies and infrastructure funds from the Fonds de Recherche en Santé Québec (FRSQ).

References


