



HAL
open science

Neural correlates of consolidation in working memory.

Nelly Mainy, Philippe Kahane, Lorella Minotti, Dominique M. Hoffmann,
Olivier Bertrand, Jean-Philippe Lachaux

► **To cite this version:**

Nelly Mainy, Philippe Kahane, Lorella Minotti, Dominique M. Hoffmann, Olivier Bertrand, et al..
Neural correlates of consolidation in working memory.. 2006. hal-00019455v1

HAL Id: hal-00019455

<https://hal.science/hal-00019455v1>

Preprint submitted on 22 Feb 2006 (v1), last revised 10 Aug 2006 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Neural correlates of consolidation in working memory.

Mainy, N. ^{1*}, Kahane, P. ², Minotti, L. ², Hoffmann, D. ², Bertrand, O. ¹, Lachaux, J.P. ¹.

¹ INSERM, U280, Lyon, F-69500, France; Institut Fédératif des Neurosciences, Lyon, F-69000, France; Univ Lyon 1, Lyon, F-69000, France

² Grenoble Hospital, Department of Neurology, Grenoble, F-38000, France ; INSERM, U318, Grenoble, F-38000, France

*Corresponding Author:

Nelly MAINY

INSERM Unit 280

69675 Bron cedex

FRANCE

Tel: 04.72.13.89.00

Fax: 04.72.13.89.01

E-mail: mainy@lyon.inserm.fr

Abstract

Many of our daily activities rely on a brain system called working memory, which implements our ability to encode information for short-term maintenance, possible manipulation and retrieval. A recent intracranial study of patients performing a paradigmatic working memory task revealed that the maintenance of information involves a distributed network of oscillations in the gamma band (>40 Hz). Using a similar task, we focused on the encoding stage, and targeted a process referred to as short-term consolidation, which corresponds to the encoding of novel items in working memory. The paradigm was designed to manipulate the subjects' intention to encode: series of ten letters were presented, among which only five had to be remembered, as indicated by visual cues preceding, or following, each letter. During this task, we recorded the intracerebral EEG of nine epileptic patients implanted in mesiotemporal structures, perisylvian regions and prefrontal areas and used time-frequency analysis to search for neural activities simultaneous with the encoding of the letters into working memory. We found such activities in the form of increases of gamma band activity in a set of regions associated with the phonological loop, including the Broca area and the auditory cortex, and in the prefrontal cortex, the pre- and post-central gyri, the hippocampus and the fusiform gyrus.

Keywords

Working memory, encoding, gamma band, intracranial EEG

Introduction

Our ability to store verbal information in the short term is a fundamental element of our everyday life, referred to as verbal working memory (Baddeley 1992). Our understanding of the fundamental mechanisms underlying this function depends on precise spatial and temporal studies of the multiple networks underlying the perception and encoding of verbal information for short-term maintenance and retrieval. Numerous functional imaging studies have identified key brain regions of those networks (Smith and Jonides 1998), but their dynamics of activation and the neural activity patterns that support their functions are less well understood.

Those last aspects have been addressed recently for the first time in two studies using local electrophysiological recordings in humans (Howard, et al. 2003; Raghavachari, et al. 2001). They recorded the intracranial EEG of epileptic patients while they learned a series of letters presented sequentially (the classic Sternberg paradigm) and found that the memory processes were accompanied with widespread energy increases in the theta (4-7 Hz) and gamma band (> 40 Hz) EEG. One striking result was that the energy measured in the gamma band increased with the memory load, that is, the number of letters already presented and stored in a series.

Those studies extended the conclusions of a large body of literature emphasizing the role of focal gamma oscillations in perception and visual short-term memory (Singer 1999; Tallon-Baudry and Bertrand 1999) to support the view that gamma oscillations participate in the perceptual and maintenance stages of working memory. A natural question is whether gamma band activities also play a role in the encoding stage of

working memory. If this is the case, then a second question is whether this network can be distinguished from the network of gamma activities mediating perception. Those are the two questions addressed by the present study.

Before reaching those questions, it is important to provide a more specific definition of the type of encoding addressed in this study, Jolicoeur and Dell'Acqua (Jolicoeur and Dell'Acqua 1998) addressed specifically the notion of encoding and distinguished between three types of encoding: sensory encoding, perceptual encoding, and short-term consolidation. In that paper, sensory encoding was defined as automatic, not subject to interference by other cognitive processes, and as corresponding to the extraction of the sensory features of the stimulus, which can give rise to sensory persistence. Perceptual encoding was defined as the process by which patterns are recognized, and the representation at this stage of encoding contains information about the identity of the patterns (such as the letters) but needs to receive bottom-up input from the sensory systems, otherwise it decays rapidly. Finally, Short-Term consolidation is defined as the process of encoding information into short-term memory, which is required for verbal report. This work echoed the proposition of Cowan ((Cowan 2001), and discussion in (Ruchkin, et al. 2003)) that visual working memory consists of at least three stages: an initial, high capacity, very limited duration iconic store that encodes the physical features of stimuli, an intermediate transient stage, and a limited capacity, postcategorical sustained short-term store. A similar view has also been expressed by (Potter 1976) and others. Taking into account these distinctions, the type of encoding studied here corresponds to the process of short-term consolidation, that is the encoding in working memory (and the term 'encoding' will be used in that specific sense in the rest of the paper).

The main difficulty is to isolate the encoding stage from the perceptual stage, because encoding usually starts as soon as an item is perceived. One way to decouple the two is to introduce a delay between the presentation of the item and the instruction to memorize it, provided that certain items must be ignored (Reber, et al. 2002).

We introduced this manipulation in the task used by (Howard, et al. 2003) to test for the existence of a network of gamma band activations associated with the encoding phase. In this paradigm, performed by nine epileptic patients who underwent intracerebral recordings, letters were presented one by one and followed (in one experimental condition) or preceded (in a second condition) by an instruction to memorize it or not. We assumed that the active encoding phase started after the presentation of the second stimulus of each instruction-letter pair, and thus shifted between the two conditions from the presentation of the instruction to the presentation of the letter. We found a network of focal gamma activations showing a similar latency shift and distributed in focal centres in the prefrontal cortex (BA 9, 10, 44, 45, 46), the precentral and postcentral gyri, the auditory cortex, the fusiform gyrus and the hippocampus.

Materials and Methods

Subjects

The 9 patients (P1 to P9) suffered from drug-resistant partial epilepsy and were candidates for surgery. There were 5 females and 4 males, aged from 21 to 53 yrs (mean: 35.7 years). Magnetic resonance imaging of the brain (MRI) showed different types of lesions in 7 cases (hippocampal sclerosis in P1-P3-P8, prefrontal dysembryoplastic neuroepithelial tumor in P7, anterolateral temporal cavernoma in P2, fronto-orbital arachnoidian cyst in P5, periventricular nodular heterotopia in P4), and was normal in the remaining 2 patients (P6-P9). Because the location of the epileptic focus could not be identified using non-invasive methods, the patients underwent intracerebral EEG recordings by means of stereotactically implanted multilead depth electrodes (SEEG) (for explanation of this methodology, see ((Kahane, et al. 2004))), on the basis of which the epileptogenic zone proved to be right temporal in 2 cases (P1-P8), left temporal in 2 cases (P2-P3), right prefrontal in 1 case (P7), left insular in 1 case (P5), and right insulo-temporal in 1 case (P9). In the remaining 2 patients, seizures proved to arise from the heterotopia in 1 (P4), and were not clearly localised in 1 (P6).

Selection of sites to implant were made entirely for clinical purposes with no reference to the present experimental protocol, however patients who entered this protocol were selected because their implantation sampled regions classically associated with verbal working memory. The patients performed the task 4 days after the implantation of the electrodes, and all had previously given their informed consent to participate in the experiment.

Electrodes implantation

11 to 14 semi-rigid electrodes were implanted per patient, in cortical areas which varied depending on the suspected origin of seizures (Fig. 1). A total number of 110 electrodes were implanted in the 9 patients, 61 in the right hemisphere and 49 in the left. Each electrode had a diameter of 0.8 mm and comprised 10 or 15 leads of 2 mm length, 1.5 mm apart (Dixi, Besançon, France), depending on the target region. Therefore, various mesial and lateral cortical areas were evaluated, including sulcal cortex. The electrode contacts were identified on each individual stereotactic scheme, and then anatomically localized using the proportional atlas of Talairach and Tournoux (Talairach and Tournoux 1988). In addition, the computer-assisted matching of post-implantation CT-scan with a pre-implantation 3-D MRI provided a direct visualization of the electrode contacts with respect to the brain anatomy of each patient (Activis, Lyon, France).

Paradigm

The paradigm was a modified version of the task used by (Howard, et al. 2003). Subjects were presented with series of ten different letters (randomly chosen between A and O), interleaved with ten central dots, either green (five dots) or red (the remaining five dots) (see Fig.1). The dots served as visual cues instructing the patients to memorize or not the previous (in a condition named AFTER) or subsequent letter (in a condition named BEFORE), and green dots served as 'yes' signal (memorize) while red dots served as 'no' signals (ignore). Within one experimental block, the patients were presented with ten such series of ten letters and ten dots and the condition was kept constant. All series followed the same temporal organization : a white fixation cross would appear for one second at the

centre of the blue computer screen, replaced by a central X-letter that stayed on screen for three seconds and served both as a fixation point and as an instruction to concentrate. The ten letters and ten cues would then alternate, each item being presented for one second, immediately followed by the next item. Four seconds after the onset of the last letter, the patients would receive a visual instruction ('reponse?') to spell aloud the five letters associated with a 'yes'-cue. Note that this is an important difference with the paradigm used by (Howard, et al. 2003), introduced to check the actual memory performance of the patients for each series of letters. A new series would then begin a couple of seconds later after a short relaxation period. The protocol included six BEFORE blocks and six AFTER blocks, intermixed, between which the patients could relax and was informed of the nature of the upcoming block (BEFORE or AFTER), and of an estimation of their performance in the last block. Subjects were explicitly instructed to remain silent during the blocks, except when they gave their responses.

All stimuli were shown foveally on a 17" computer screen. The stimulus dimensions were 4 cm x 4 cm for the letters and 2 cm x 2 cm for the dots, subtending a visual angle of approximately 1 x 1 degrees and 0.5 x 0.5 degrees at the 200 cm viewing distance.

Recordings

Intracerebral recordings were conducted using an audio-video-EEG monitoring system (Micromed, Treviso, Italy), which allowed the simultaneous recording of 63 depth-EEG channels sampled at 512 Hz [0.1–200 Hz bandwidth] during the

experimental paradigm. One of the contact sites in the white matter was chosen as reference. This reference has the same impedance as the other contact sites, and was located in a region with no or little source of electrical field, in addition, it was not contaminated by eye-movements artifacts or electromyographic activity from subtle muscle contractions. However, all signals were re-referenced to their nearest neighbour on the same electrode, 3.5 millimeters away before analysis (bipolar montage). Recording sites showing clear epileptiform activities were excluded from the analysis, and among the remaining sites, monopolar and bipolar data were systematically inspected, both raw and high-pass filtered (above 15 Hz), and any trial showing epileptic spikes in any of those traces was discarded.

Time-frequency analysis

For each single trial, bipolar derivations computed between adjacent electrode contacts were analyzed in the time-frequency domain by convolution with complex Gaussian Morlet's wavelets (Tallon-Baudry, et al. 1996), thus providing a time-

frequency power map $P(t, f) = |w(t, f) * s(t)|^2$, where $w(t, f)$ was for each time t and

frequency f a complex Morlet's wavelet $w(t, f) = A \exp(-t^2 / 2\sigma_t^2) \cdot \exp(2i\pi ft)$, with

$A = (\sigma_t \sqrt{\pi})^{-1/2}$ and $\sigma_t = 1/(2\pi\sigma_f)$ and σ_f a function of the frequency f : $\sigma_f = f / 7$.

Normalized time-frequency plots were computed for each bipolar derivation, for visualization purpose. This normalization was done separately for each frequency, and consisted in a) subtracting the mean power during a [-500ms : -100 ms]

prestimulus baseline, and b) dividing by the standard deviation of the power during this same baseline.

Comparison between pairs of cue/letter combinations (e.g. 'yes'-cue/letter vs 'no'/letter) were done via a Kruskal-Wallis non-parametric analysis (chosen because it makes no assumption about the distribution of the data, (Kruskal and Wallis 1952; Wikipedia 2006)) applied on the raw time-frequency values of energy, on a set of time-frequency tiles [100ms x 30 Hz] covering a [0 : 2000 ms] x [5 : 250 Hz] domain. There was one test per tile comparing the values obtained for all the trials in the two conditions, in total, the [0 : 2000 ms] x [5 : 250 Hz] time-frequency domain was covered by $39 \times 14 = 546$ [100ms x 30 Hz] tiles (considering an overlap of 50% along the time and frequency dimensions). To apply a Bonferroni correction, we set the threshold for significance for each individual test to $p = 0.0001$, which yields a corrected threshold of 0.05 for all the tests. Note that this Bonferroni correction, which assumes that all the tests are independent, was certainly too conservative, since the energy values measured in consecutive tiles were not independent because of the limited time and frequency resolution of the wavelet transform. However, there is no simple manner to perform an exact correction as the correlation between consecutive energy values depends on frequency, as does the temporal and frequency resolution of the wavelet transform.

Electrodes presented in the result section were selected because they reached the criterion of $p < 0.0001$ in at least one of the time-frequency tiles. However, we did not consider in this study isolated effects, that is, effects found in one brain region in only one patient, but with no equivalent in the other patients. This strategy was to compensate the limitation imposed by the fact that in a given patient, a specific brain

region is sampled by only one electrode: therefore, since the reproducibility of an effect cannot be tested across sites in a single patient, we impose that it is compensated by a reproducibility in at least another patient with an electrode in the same functional structure.

Note that EEG signals were evaluated with the software package for electrophysiological analysis (ELAN-Pack) developed in the INSERM U280 laboratory.

Results

The patients were successful at memorizing the letters when instructed to do so by the cue (the recall percentage of 91.6 % +/- 6.6 % for the letters associated with a memorization instruction). We could not test systematically whether the other letters were memorized or not, because any question concerning those 'ignored' letters would have brought the patient's attention to them. However, when we occasionally asked the patients whether they could remember some of the non-target letters their answer was either that they did not pay attention to them, or that our question was weird since they were instructed to ignore them.

The data analysis compared for each recorded site the energy modulations induced in the gamma band by 'yes' vs 'no' cue/letter sequences in a search for cue-related modulations defined as significant energy differences between the two conditions

(Kruskal-Wallis test, $p < 0.0001$, see Methods); this was done separately in the BEFORE and the AFTER condition.

Figure 2 illustrates the key manipulation of this paradigm. In the BEFORE condition, the cues instructing the patients to memorize or not the associated letters come before the letters, while in the AFTER condition, they come after the letters. Therefore, if we assume that the process of actively encoding a target letter into working memory should start only after both the letter and the memorization instruction (the cue) have been presented (see black frame in Fig 2c), then cue-related gamma energy modulations should follow this shift and switch between the two conditions from the letter onsets (in the BEFORE condition) to the cue onsets (in the AFTER condition), as they should if associated with the encoding process.

The time frequency statistical analysis revealed cue-related energy modulations in the gamma band in a subset of the recording sites (17 out of 324) (by cue-related modulation, we mean an energy modulation in response to a cue or to a letter which varies with the value of the cue). Those were confined in small anatomical clusters located in the posterior aspect of the prefrontal cortex (BA 44, 46), which corresponds, in the dominant hemisphere for language, to the pars opercularis of Broca area), the precentral and the postcentral gyrus, the auditory cortex, the prefrontal cortex, the fusiform gyrus and the hippocampus (see table I for the list of the activated sites with their Talairach coordinates). The effects were observed both in the BEFORE and the AFTER conditions: however, the latencies of the effects varied across conditions. In the BEFORE condition, the effect was predominantly after the letter: letters preceded by a 'yes' cue induced an increase in gamma energy

while those preceded by a 'no' cue did not (except in the hippocampus, where no increase was observed in this condition). The statistical analysis revealed that the effect typically extended over a wide (40 Hz – 150 Hz) frequency band and lasted up to several hundreds of milliseconds after the letter presentation, with latencies variable across sites. In the AFTER condition, the effect was observed after the cue presentation and before the letter presentation: 'yes' cues induced a gamma energy increase while 'no' cues did not. Thus, the cue-related gamma energy modulations, when present, followed the predicted latency shift between conditions: as illustrated in Figure 2c, the modulation shifted between the BEFORE condition and the AFTER condition, from the onset of the letter to the onset of the cue, one second earlier.

Figures 2 to 6 illustrate those gamma energy modulations for each anatomical cluster, with the recording sites reconstructed on the MNI single subject MRI. In each figure, the displays are organized so that the gamma band responses in the four conditions (responses to target and non-target letters in the BEFORE and AFTER conditions) can be compared, in amplitude and in timing. As mentioned above, the effects most clearly associated with encoding are the responses which latency is shifted between the BEFORE (i.e. left column) and AFTER (i.e. right column) conditions. Figure 2 shows the two sites where a cue-related modulation was observed in the Broca area pars Opercularis in the left hemisphere, and the time-frequency energy variations across conditions and memorization instruction for one of those sites. The sharp gamma band energy increase is clearly visible in both conditions, and shifts from the letter onset to the cue onset between the BEFORE and the AFTER condition. Similar effects were observed in two sites in the pre-central gyrus and two sites in the post-central gyrus, i.e. in regions associated

respectively with the motor and sensory representation of the face and larynx (Fig 3). In the auditory cortex (Fig 4), the shift was observed in four sites from four different patients, two in each hemisphere, and interestingly the effect started slightly later than in the Broca area (compare Fig 5 and 2). The other sites with an effect were in the Prefrontal Cortex (2 sites, Fig 5), in the fusiform gyrus (2 sites, Fig 6) and in the hippocampus. The two sites in the fusiform gyrus were located in symmetrical positions in the two hemispheres, in a region close to the proposed Word Form Area (Cohen, et al. 2000) (see table I for Talairach coordinates). In this region of the visual ventral pathway, gamma energy bursts occurred not only in response to the letters but also in response to the cues, and in the BEFORE condition, the energy was stronger for the 'yes'-cues, in addition to the cue-related modulation of the gamma peak induced by the letter, although the later effect was broader. Finally, the effect in the hippocampus was atypical, different from the other sites, since cue-related gamma modulations were only observed in the AFTER condition.

Discussion

Working memory involves at least four stages: perceptual and memory encoding, maintenance and retrieval. Those stages are usually tightly bound together, and most of the studies that have tried to identify their individual neural substrates have acknowledged the difficulty to decouple them (Manoach, et al. 2003). The purpose of the present study was to single out the memory encoding stage and to correlate it with precisely timed neural events. We believe that we found such correlates in the form of a distributed network of focal gamma band activations, part of which matches relatively well to what has been described as the phonological loop.

Several groups have successfully used fMRI or PET to identify key brain structures for each of those four memory stages, through smart experimental designs and analysis techniques (Manoach, et al. 2003; Reber, et al. 2002) that managed to decouple them. However, those functional imaging techniques could not resolve the flow of information processing that occurs over time scales on the order of tens of milliseconds (Buckner and Koutstaal 1998) and they must ultimately be completed by focal electrophysiological studies of those brain structures, as available only in epileptic patients. Such recordings are necessary a) to measure directly the absolute and relative latencies of the neural activations within those structures b) to follow the quick transition between memory stages and c) to associate those stages with actual neural processes (Lachaux, et al. 2003).

One of the first studies to investigate verbal working memory at such a fine resolution revealed gamma oscillations which amplitude covaried with memory load, which is a

property characteristic of a neural correlate of the maintenance process (Howard, et al. 2003). The present study shifted the focus of investigation to the encoding phase of verbal working memory, which neural substrates remain unclear. Our framework draws on recent observations (Fell, et al. 2001; Sederberg, et al. 2003) showing that the amplitude of the gamma oscillations induced by the presentation of words were predictive of their successful encoding into long-term memory. Although long-term memory and working memory are distinct processes, we hypothesized that the same would be true for verbal working memory.

Our paradigm was based on the logical premise that in average encoding starts when a subject has a) an item to encode and b) the intention to encode it. Note once again that the term 'encoding' refers here to the consolidation in working memory, as presented in the introduction in reference to (Jolicoeur and Dell'Acqua 1998) (sensory and perceptual encoding certainly starts as soon as a letter is presented in both the AFTER and BEFORE conditions). Also, we do not claim that all the target letters are encoded while none of the non-target letters are, since in some instances items that are not to be encoded (non-target letters) can still be, in a form of incidental encoding. What we claim is that our experimental paradigm created a bias between the two categories of letters, and that the target letters yielded in average a deeper consolidation process than the non-target letters. This last claim is based on the repeated observation that the memory span for letters is typically below 7 ((Miller 1956), (Cowan 2001)): since the subjects did recollect more than 4 out of the 5 target letters in average (see the results section), it is unlikely that the non-target letters were memorized as well, because this would have required from the subjects an average memory span above 8.

By presenting first the item or the instruction, we could therefore move in time the moment of the encoding, and we found gamma energy modulations which followed this manipulation. Those gamma modulations may therefore constitute part of the neural substrate of this encoding process. Note that those gamma modulations extended in a broad gamma range, consistent with previous intracranial observations from our group and others (Lachaux, et al. 2005; Tanji, et al. 2005).

Before going further, we should bring a couple of qualifications to our strong conclusions. First, we are fully aware, though, that those results are correlative by nature, as in any functional imaging study and do not constitute an indisputable proof that encoding does require this network of gamma activations. Then, we should keep in mind that the sampling of brain tissue provided by those implanted electrodes was limited to a small portion of the brain, such that important functional areas involved in memory encoding may not have been explored. Also, we do not claim that those activations are the only neural substrate of the encoding phase, even in this region, as low frequency evoked potentials may also participate in this function. Further, we cannot exclude the possibility that at least in some cases, the difference in memorization between target and non-target letters occurs not at the level of the encoding, but at a subsequent stage which would include an active forgetting of the non-targets, a phenomenon demonstrated by (Anderson, et al. 2004). However, figure 7 which shows the time course of gamma activity in a frontal site when the five non-target letters are presented first seems to contradict this possibility: the graph shows clearly that the change in neural activity occurs solely in response to the target letters, and not in response to the non-target letters as would be expected if it was

related to a suppressive mechanism. Finally, although our study focused on the encoding phase, we also observed gamma band activations that increased in power with the memory load (see fig 7). This confirmed the link established by (Howard, et al. 2003) between gamma oscillations and the maintenance phase. Together, our two studies support the proposition that gamma activity serves to organize the content of working memory (Jensen and Lisman 1998).

Nevertheless, our study was to our knowledge the first effort to find correlations between the encoding phase of verbal working memory and direct measures of local neural activities. It is therefore impossible to compare our findings with observations of similar nature, however, it is tempting to try a comparison with previous functional imaging studies, provided a spatial congruence between the neural phenomena quantified by metabolic measures and gamma band activations (a correspondence between the BOLD signal and gamma band activities has received recent support (Logothetis 2002). Several of the gamma activations observed in this study were located in brain regions classically associated with speech, such as the Broca area (pars opercularis) and the pre and post central sensory-motor representations of the face and larynx. Indeed, all those regions showed a strong gamma activation when the patients gave their verbal responses, at the end of each series of letters (Fig 7).

More precisely, the pars opercularis of Broca has been associated with the manipulation of phonology (Fiez 1997; Jobard, et al. 2003), with subvocal rehearsal (Paulesu, et al. 1993) and with verbal working memory (Fiez, et al. 1996). The precentral and postcentral gamma activations can be interpreted as sensori-motor activations during articulatory motor sequences subtending subvocalisations, as identified by functional imaging (Fox, et al. 1987; Perry, et al. 1999). This strongly

suggest that implicit speech has been operative in those patients during the encoding of the letters. This is further confirmed by the fact that gamma activations were observed in the auditory cortex, a structure also activated by subvocalization (Belin, et al. 2000; Smith, et al. 1995) and auditory imagery (Kraemer, et al. 2005). Also, introspective reports from the subjects suggested that they indeed spelled out the letters mentally.

This anatomical organization is partly consistent with the model of the phonological loop, introduced by Baddeley in its influential model of working memory (Baddeley 2003b) as 'a phonological store, which can hold memory traces for a few seconds before they fade out, and an articulatory rehearsal process that is analogous to subvocal speech' (Baddeley 2003a). The functions performed by the phonological loop are explicitly those necessary to perform the Sternberg task and our results confirm that the loop may also participate in the encoding of novel items, possibly via an articulatory process that would update the content of the memory store by adding the new item to the list already stored. Our results also confirm that this network is not solely devoted to a continuous articulatory rehearsal, as correctly suggested by (Cohen, et al. 1997).

Our observations are in-line with the previous fMRI studies which have localized components of the phonological loop in the frontal speech areas (Smith and Jonides 1998). It is in fact quite remarkable, given that the exact relationship between BOLD and gamma oscillations is still unknown, that the two phenomena tend to gather in the same regions. Note that many fMRI studies have also found activations in the posterior parietal cortex when short-term storage of verbal material was needed (Smith and Jonides 1998), but our electrodes did not sample this region.

Other gamma effects were observed in the prefrontal cortex, the hippocampus and the fusiform gyrus and are unlikely to mediate directly the phonological loop. The prefrontal cortex has been frequently found to be active during working memory tasks (Cohen, et al. 1994; Miller, et al. 1996), but in the Dorsal Lateral Prefrontal Cortex, which could mediate the manipulation of stored information (Smith and Jonides 1998) and attention load (Mazoyer, et al. 2002). In the present study, the activation sites were more ventral, slightly anterior to Broca and therefore likely involved in implicit speech. The same is probably not true for the gamma band effects observed in the hippocampus and the fusiform gyrus. In the hippocampus, cue-related modulations were only observed in the AFTER condition, in response to the cue onset, which may be surprising in light of recent evidence that the gamma oscillations in the hippocampus play a role in the encoding phase, at least for words and in long-term memory (Fell, et al. 2001). We have no straightforward interpretation for this, one should simply note that the AFTER condition is the only one in which the encoding takes place in the absence of the item to encode (in fact, right after it disappears). In our particular protocol, the hippocampus may not be necessary when the item to encode is still physically available to the subject (as in the BEFORE condition).

In the fusiform gyrus, the cue-effect had two faces. In addition to the shift apparent in the other regions, that we have associated with the encoding phase, a cue-effect was observed in response to the dots in the BEFORE condition. This effect cannot be attributed to the encoding of the subsequent letter, and must probably be thought of as a visual attention effect, in the light of recent studies showing that visually induced gamma activations in the fusiform gyrus are very sensitive to attentional modulations particularly in the period preceding the attended stimulus (Tallon-Baudry, et al. 2005).

If this interpretation is correct, then this task would bring into play two systems of gamma activations: one for the perception of the items, and modulated by attention, the other for their encoding into working memory and modulated by the subjects intention to memorize.

This should be emphasized since it would provide a neural basis for the distinction between two major causes of fluctuations in memory performance: fluctuations in the attention to the items, and fluctuations in the intention to memorize them (Reber, et al. 2002). It is a familiar fact, validated by behavioral experiments (Kensinger, et al. 2003; Naveh-Benjamin, et al. 2003), that memory suffers from divided attention during learning. And it has been shown that this is most likely the encoding phase which is affected (Iidaka, et al. 2000). However, it is often difficult to dissociate actual perceptual attention modulations from modulations of the subject's intention to memorize (Reber, et al. 2002). Our results suggest that the two phenomena operate on two different systems of gamma activations: one mediating the perception of the verbal information and one mediating the encoding of this information. Fluctuations in the attention to the item certainly correspond to energy fluctuations in the first of those two systems, as supported by several human and animal studies (Fell, et al. 2003; Fries, et al. 2001) and as confirmed by our observations in the fusiform gyrus (Lachaux, et al. 2005). We provide evidence that the second of those two networks is sensitive to variations in the subject's intention to encode, which can thus be measured quantitatively. Therefore, the monitoring of gamma activity in specific, distributed brain regions may constitute an objective measure of encoding intention and perceptual attention fluctuations during working memory paradigms.

Acknowledgements

We thank Valérie Balle, Patricia Boschetti, Carole Chatelard, Véronique Dorlin, Eliane Gamblin and Martine Juillard for their invaluable help. JPL was supported by the Fondation Fyssen.

Figure Legends

Figure 1: Experimental paradigm and recording sites. a) entry points of the intracranial electrodes across all patients, represented on a 3D reconstruction of the MNI single subject MRI. The electrodes are rods oriented perpendicular to the plane of the figure. A total number of 110 electrodes were implanted, 49 in the left hemisphere (left side of the figure), and 61 in the right hemisphere (right side of the figure) b) experimental paradigm. In the BEFORE condition, the subjects had to memorize the five-out-of-ten letters preceded by green dots (the 'yes'-cues); in the AFTER condition, the subjects had to memorize the five-out-of-ten letters followed by a green dot. Red dots meant for the patient that s/he did not have to memorize the associated letter. Each item (dot or letter) stayed on screen for one second and was immediately replaced by the next item. Memory encoding is expected to start only when the subject has an item to memorize and has been instructed to do so (some degree of automatic encoding may start during the letter presentation in the AFTER condition, but the task does not encourage this, because the letter stays onscreen until the subsequent cue appears). The moment of encoding should therefore shift from the letter onset in the BEFORE condition, to the green-dot/'yes'-cue onset in the AFTER condition. Red arrows point towards the x,y and z directions in the Talairach space.

Figure 2: A possible neural correlate of memory encoding in Broca area (pars opercularis). a) time-frequency representation showing the EEG energy modulation induced in Broca area by 'yes'-cue/letter pairs in the BEFORE (Bg) and AFTER (Ag) conditions (each time-course is expressed in units of the standard deviation of the [-500 ms : -100 ms] prestimulus period); In all maps, the cue occurs at 0 ms and the

subsequent letter at 1000 ms. (Br) and (Ar) : same for 'no'-cue/letter pairs. b) statistical time-frequency maps showing the p-values of the Kruskal-Wallis comparison between the 'yes' cue/letter and 'no' cue/letter pairs in the BEFORE (SB) and AFTER (SA) conditions, what we called 'cue-related effects' correspond to time-frequency areas with $p < 0.001$. c) Schematic prediction of the moment when memory encoding should start (black frame) in the BEFORE (left) and AFTER (right) conditions. d) Time profiles of the average energy in the [40-150 Hz] band for each of the maps shown in a) (Ag, Ar, Bg, Br). e) and f) are reconstructions on the MNI single subject MRI of the recording positions in Broca area with cue-related effects in the gamma band (P3, r'8 site, the illustrated example and P2, r'6 site. See table I for their Talairach coordinates). All time-frequency maps correspond to the EEG recorded in r'8, the site indicated by the black arrow.

Figure 3: same as figure 2 for the precentral and postcentral recording sites. s10 (in P8) is the example, the other sites are s'8 (P3), h'12 (P5) and h14 (P9) (see table I for Talairach coordinates).

Figure 4: same as figure 2 for the recordings sites in the auditory cortex. w3 (in P7) is the example shown, the other sites being x2 (P1), u'3 (P2) and u'7 (P3).

Figure 5: same as figure 2 for the prefrontal sites. g'12 (in P6) is the example, the other site is q2 (P9)

Figure 6: same as figure 2 for the sites in the fusiform gyrus. e'5 (in P1) is the example, the other sites being f'8 (P2) and e'9 (P4).

Figure 7: Gamma Band Energy correlates with memory load in Broca Area. We used an additional control condition, similar as BEFORE, except that the patients knew in advance that the five letters to memorize would always be the last ones to be presented. This figure shows the average energy in the [50 -150 Hz] band across the entire sequence of letters for a recording site located in Broca area (the same site chosen to illustrate figure 2). The energy shows a steady increase throughout the five letters to memorize, in addition to rapid bursts at the onset of the letters. Note also the strong increase at the time of the subject's response.

Table 1: Talairach coordinates for the recording sites with cue-related gamma energy modulations. Each brain region is the object of a specific figure (2 to 6) showing the corresponding recording sites (except for the hippocampus). The coordinates of the sites chosen as illustrative examples in each figure are written in bold font. X,y and z refer to the Talairach coordinates of the sites, in millimeters (not the MNI coordinates). Those values were converted to MNI coordinates only for the purpose of representing the sites onto the MNI single subject MRI.

Patient	Anatomical location	L/R	Talairach coordinates		
			x	y	z
	BROCA				
2	R'6	L	-59	2.1	14
3	R'8	L	-65	5.8	12
	PRE/POST-CENTRAL				
5	H'12	L	-55	-4.3	37
3	S'8	L	-69	-15	24
9	H14	R	64	-19	31
8	S10	R	65	-19	19
	AUDITORY CORTEX				
3	U'7	L	-62	-26	7.8
2	U'3	L	-53	-30	6.1
7	W3	R	39	-18	11
1	X2	R	32	-22	10
	PREFRONTAL				
9	Q2	R	33	13	8.3
6	G'12	L	-54	39	11
	FUSIFORM GYRUS				
4	E'9	L	-51	-51	-12
1	E'5	L	-42	-42	-12
2	F'8	L	-50	-52	-12
	HIPPOCAMPUS				
3	E'2	L	-27	-17	-21
1	B'3	L	-31	-27	-9.3

REFERENCES

- Anderson MC, Ochsner KN, Kuhl B, Cooper J, Robertson E, Gabrieli SW, Glover GH, Gabrieli JD. (2004): Neural systems underlying the suppression of unwanted memories. *Science* 303(5655):232-5.
- Baddeley A. (1992): Working memory. *Science* 255(5044):556-9.
- Baddeley A. (2003a): Working memory and language: an overview. *J Commun Disord* 36(3):189-208.
- Baddeley A. (2003b): Working memory: looking back and looking forward. *Nat Rev Neurosci* 4(10):829-39.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. (2000): Voice-selective areas in human auditory cortex. *Nature* 403(6767):309-12.
- Buckner RL, Koutstaal W. (1998): Functional neuroimaging studies of encoding, priming, and explicit memory retrieval. *Proc Natl Acad Sci U S A* 95(3):891-8.
- Cohen JD, Forman SD, Braver TS, Casey BJ, Servan-Schreiber D, Noll DC. (1994): Activation of prefrontal cortex in a non-spatial working memory task with functional MRI. *Human Brain Mapping* 1:293-304.
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE. (1997): Temporal dynamics of brain activation during a working memory task. *Nature* 386(6625):604-8.
- Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, Henaff MA, Michel F. (2000): The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123 (Pt 2):291-307.
- Cowan N. (2001): The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci* 24(1):87-114; discussion 114-85.
- Fell J, Fernandez G, Klaver P, Elger CE, Fries P. (2003): Is synchronized neuronal gamma activity relevant for selective attention? *Brain Res Brain Res Rev* 42(3):265-72.
- Fell J, Klaver P, Lehnertz K, Grunwald T, Schaller C, Elger CE, Fernandez G. (2001): Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nat Neurosci* 4(12):1259-64.
- Fiez JA. (1997): Phonology, semantics, and the role of the left inferior prefrontal cortex. *Hum Brain Mapp* 5(2):79-83.
- Fiez JA, Raife EA, Balota DA, Schwarz JP, Raichle ME, Petersen SE. (1996): A positron emission tomography study of the short-term maintenance of verbal information. *J Neurosci* 16(2):808-22.
- Fox PT, Burton H, Raichle ME. (1987): Mapping human somatosensory cortex with positron emission tomography. *J Neurosurg* 67(1):34-43.
- Fries P, Reynolds JH, Rorie AE, Desimone R. (2001): Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291(5508):1560-3.
- Howard MW, Rizzuto DS, Caplan JB, Madsen JR, Lisman J, Aschenbrenner-Scheibe R, Schulze-Bonhage A, Kahana MJ. (2003): Gamma oscillations correlate with working memory load in humans. *Cereb Cortex* 13(12):1369-74.

- Iidaka T, Anderson ND, Kapur S, Cabeza R, Craik FI. (2000): The effect of divided attention on encoding and retrieval in episodic memory revealed by positron emission tomography. *J Cogn Neurosci* 12(2):267-80.
- Jensen O, Lisman JE. (1998): An oscillatory short-term memory buffer model can account for data on the Sternberg task. *J Neurosci* 18(24):10688-99.
- Jobard G, Crivello F, Tzourio-Mazoyer N. (2003): Evaluation of the dual route theory of reading: a metaanalysis of 35 neuroimaging studies. *Neuroimage* 20(2):693-712.
- Jolicoeur P, Dell'Acqua R. (1998): The demonstration of short-term consolidation. *Cognit Psychol* 36(2):138-202.
- Kahane P, Minotti L, Hoffmann D, Lachaux J, Ryvlin P. 2004. Invasive EEG in the definition of the seizure onset zone: depth electrodes. In: Rosenow F, Lüders HO, editors. *Handbook of Clinical Neurophysiology. Pre-surgical assessment of the epilepsies with clinical neurophysiology and functional neuroimaging*: Elsevier Science.
- Kensinger EA, Clarke RJ, Corkin S. (2003): What neural correlates underlie successful encoding and retrieval? A functional magnetic resonance imaging study using a divided attention paradigm. *J Neurosci* 23(6):2407-15.
- Kraemer DJ, Macrae CN, Green AE, Kelley WM. (2005): Musical imagery: sound of silence activates auditory cortex. *Nature* 434(7030):158.
- Kruskal WH, Wallis WA. (1952): Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47(260):583–621.
- Lachaux JP, George N, Tallon-Baudry C, Martinerie J, Hugueville L, Minotti L, Kahane P, Renault B. (2005): The many faces of the gamma band response to complex visual stimuli. *Neuroimage* 25(2):491-501.
- Lachaux JP, Rudrauf D, Kahane P. (2003): Intracranial EEG and human brain mapping. *J Physiol Paris* 97(4-6):613-628.
- Logothetis NK. (2002): The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos Trans R Soc Lond B Biol Sci* 357(1424):1003-37.
- Manoach DS, Greve DN, Lindgren KA, Dale AM. (2003): Identifying regional activity associated with temporally separated components of working memory using event-related functional MRI. *Neuroimage* 20(3):1670-84.
- Mazoyer P, Wicker B, Fonlupt P. (2002): A neural network elicited by parametric manipulation of the attention load. *Neuroreport* 13(17):2331-4.
- Miller EK, Erickson CA, Desimone R. (1996): Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16(16):5154-67.
- Miller GA. (1956): The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 63(2):81-97.
- Naveh-Benjamin M, Guez J, Marom M. (2003): The effects of divided attention at encoding on item and associative memory. *Mem Cognit* 31(7):1021-35.
- Paulesu E, Frith CD, Frackowiak RS. (1993): The neural correlates of the verbal component of working memory. *Nature* 362(6418):342-5.
- Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC. (1999): Localization of cerebral activity during simple singing. *Neuroreport* 10(18):3979-84.
- Potter MC. (1976): Short-term conceptual memory for pictures. *J Exp Psychol [Hum Learn]* 2(5):509-22.
- Raghavachari S, Kahana MJ, Rizzuto DS, Caplan JB, Kirschen MP, Bourgeois B, Madsen JR, Lisman JE. (2001): Gating of human theta oscillations by a working memory task. *J Neurosci* 21(9):3175-83.

- Reber PJ, Siwiec RM, Gitelman DR, Parrish TB, Mesulam MM, Paller KA. (2002): Neural correlates of successful encoding identified using functional magnetic resonance imaging. *J Neurosci* 22(21):9541-8.
- Ruchkin DS, Grafman J, Cameron K, Berndt RS. (2003): Working memory retention systems: a state of activated long-term memory. *Behav Brain Sci* 26(6):709-28; discussion 728-77.
- Sederberg PB, Kahana MJ, Howard MW, Donner EJ, Madsen JR. (2003): Theta and gamma oscillations during encoding predict subsequent recall. *J Neurosci* 23(34):10809-14.
- Singer W. (1999): Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24(1):49-65, 111-25.
- Smith EE, Jonides J. (1998): Neuroimaging analyses of human working memory. *Proc Natl Acad Sci U S A* 95(20):12061-8.
- Smith JD, Wilson M, Reisberg D. (1995): The role of subvocalization in auditory imagery. *Neuropsychologia* 33(11):1433-54.
- Sperling G. (1960): The information available in brief visual presentations. *Psychological Monographs: General and Applied* 74(498):1-29.
- Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: an approach to cerebral imaging. New York: Thieme.
- Tallon-Baudry C, Bertrand O. (1999): Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci* 3(4):151-162.
- Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. (1996): Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci* 16(13):4240-9.
- Tallon-Baudry C, Bertrand O, Henaff MA, Isnard J, Fischer C. (2005): Attention modulates gamma-band oscillations differently in the human lateral occipital cortex and fusiform gyrus. *Cereb Cortex* 15(5):654-62.
- Tanji K, Suzuki K, Delorme A, Shamoto H, Nakasato N. (2005): High-frequency gamma-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. *J Neurosci* 25(13):3287-93.
- Wikipedia. 2006. Kruskal-Wallis one-way analysis of variance.