University of Birmingham
Research at Birmingham

Reactivation of neural patterns during memory reinstatement supports encoding specificity

Hanslmayr, Simon

DOI:
10.1080/17588928.2019.1621825

License:
None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
Hanslmayr, S 2019, 'Reactivation of neural patterns during memory reinstatement supports encoding specificity', *Cognitive Neuroscience*, vol. 10, no. 4, pp. 175-185. https://doi.org/10.1080/17588928.2019.1621825

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 03/07/2019
This is an Accepted Manuscript of an article published by Taylor & Francis in Cognitive Neuroscience on 12/06/2019, available online: https://doi.org/10.1080/17588928.2019.1621825

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- •Users may freely distribute the URL that is used to identify this publication.
- •Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- •User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
- •Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 19. Apr. 2024

Reactivation of Neural Patterns during Memory Reinstatement supports Encoding Specificity

Tobias Staudigl[1,*], Simon Hanslmayr[2,*]

1. Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, USA

2. University of Birmingham, School of Psychology; Birmingham, United Kingdom

* Address correspondence to Tobias Staudigl (tobias.staudigl@cshs.org), or Simon Hanslmayr (s.hanslmayr@bham.ac.uk)

Author Contributions

T.S. and S.H. designed the experiment and wrote the paper. T.S. collected the data. T.S. performed the analyses, S.H. performed the simulation

Competing Financial Interest

The authors declare no competing financial interests.

Abstract

Encoding specificity states that encoding and retrieving items in the same modality benefits memory, compared to encoding and retrieving in different modalities. In neural terms, this can be expressed as memory cues resonating with stored engrams; the more they overlap the better memory performance. We used temporal pattern analysis in MEG in a sensory match/mismatch memory paradigm (i.e. items presented aurally or visually) to track this resonance process. A computational model predicted that reactivation of encoding-related sensory patterns has opposing effects depending on the match or mismatch between memory cue and encoding modality. Behavioural performance was better in the match than the mismatch condition. Neural pattern reinstatement of MEG activity benefitted memory only in the match condition, but impaired memory in the mismatch condition. These effects were only obtained for aurally but not visually encoded words. The results suggest that reactivation of encoding-related neural patterns underlies encoding specificity.

Introduction

A memory cue is more effective when it overlaps with what has been studied. For instance, we are doing better in remembering when the context between encoding and retrieval is the same, compared to when they differ. A classic demonstration of this effect comes from Godden and Baddeley (Godden and Baddeley, 1975) who showed that the chances of retrieving a memory are higher when the environmental context during retrieval is the same as the one in which the item has been studied; i.e. items that are studied and retrieved on land (as opposed to when they are studied under water and retrieved on land). Similar effects have been observed with matching/non-matching background movies (Smith and Manzano, 2010; Staudigl, Vollmar, Noachtar, & Hanslmayr, 2015), or with different encoding/retrieval operations (i.e. rhyme vs semantic processing, (Morris, Bransford, & Franks, 1977). Similar effects have also been found with manipulating the modality in which items are presented at encoding and retrieval, for instance using pictures and words (Bauch and Otten, 2012; Mcdermott and Roediger, 1994), or presenting words visually and aurally (Mulligan and Osborn, 2009). In cognitive psychology these match/mismatch effects can be subsumed under the terms Encoding Specificity Principle (Tulving and Thomson, 1973), or Transfer Appropriate Processing (Morris, et al., 1977). Both concepts highlight a basic organizing principle of human memory: The likelihood of remembering a certain memory increases with the degree to which a reminder that is provided during retrieval, resonates with a stored engram of that memory, i.e. interacts with a neural pattern that was established during encoding (Rugg, Johnson, Park, & Uncapher, 2008; Tulving, 1983).

Supporting evidence for a role of neural pattern reinstatement for memory retrieval comes from a number of studies showing reinstatement of encoding patterns during retrieval (Jafarpour, Fuentemilla, Horner, Penny, & Duzel, 2014; Johnson, McDuff, Rugg, & Norman, 2009; Kerrén, Linde-Domingo, Hanslmayr, & Wimber, 2018; S. Michelmann, Bowman, & Hanslmayr, 2016; Sebastian Michelmann, Bowman, & Hanslmayr, 2018; Sebastian Michelmann, Staresina, Bowman, & Hanslmayr, 2018; Polyn, Natu, Cohen, & Norman, 2005; Schreiner, Doeller, Jensen, Rasch, & Staudigl, 2018; Staresina, Henson, Kriegeskorte, & Alink, 2012; Staresina et al., 2016; Yaffe et al., 2014). These neural patterns can be detected using various multivariate analysis approaches in order to identify reactivation of neural patterns in space (Jafarpour, et al., 2014; Polyn, et al., 2005; Staresina, et al., 2012), time (S. Michelmann, et al., 2016; Staudigl, et al., 2015), and time-frequency (Staresina, et al., 2016; Yaffe, et al., 2014). These studies demonstrate the general importance of neural pattern reinstatement for memory retrieval, however, we know little about the functional relevance of neural pattern reinstatement for memory and its involvement in Encoding Specificity (or Transfer Appropriate Processing). This is because there are hardly any studies which have tested the impact of neural pattern reinstatement on memory retrieval in the face of retrieval cues which match or do not match the encoding modality.

Two predictions can be made as to how neural pattern reinstatement impacts on memory in such matching and non-matching retrieval contexts. The first prediction is that the reinstatement of the original encoding pattern during retrieval is always beneficial for memory; regardless of a match between the retrieval and study contexts. In such a scenario, reactivation of the neural pattern containing the study episode does not support Encoding Specificity. A contrasting second prediction is that the reinstatement of encoding patterns is only beneficial for retrieval when the study and retrieval contexts match; if the study and retrieval context do not match, then reinstatement of encoding patterns might even be detrimental. In this case, the reactivated encoding pattern does not match the retrieval cue. This conflict could bias the memory decision towards a miss, i.e., judging an old item as new. In this second scenario, reactivation of the neural pattern containing the study episode would support Encoding Specificity. To the best of our knowledge, only one such study has been conducted which investigated neural pattern reinstatement in a paradigm using background movies as matching/non-matching contexts

(Staudigl, et al., 2015). Indeed, in this prior study we confirmed the second prediction, showing that memory benefitted from neural pattern reinstatement only when the encoding and retrieval contexts matched but suffered from reinstatement when the contexts did not match. This study can be seen as a first evidence that neural patterns established during encoding resonate with a retrieval cue, and that this resonance has a functional relevance for memory retrieval. This result therefore serves as a neural explanation for the Encoding Specificity Principle (or Transfer Appropriate Processing). However, if this reactivation of neural patterns is a general mechanism underlying Encoding Specificity, then we should be able to expand these results to different match/mismatch manipulations. It is therefore necessary to investigate specific predictions about neural pattern reactivation in different matching and mismatching encoding and testing scenarios. Specifically, it is important to show that the interaction between neural pattern reactivation and encoding-retrieval overlap is not only observed in visual contextual overlap (as tested in Staudigl et al., 2015) but is also observed in other sensory modalities. Therefore, we here investigate neural pattern reactivation in a memory paradigm using a sensory modality match manipulation (Mulligan and Osborn, 2009) in which the items are studied and retrieved either visually or aurally (Figures 1 & 2). We demonstrate the effects of the modality match on memory performance in a simple computational model (Figure 2), illustrating our hypothesis in support of Encoding Specificity.

Neural pattern reinstatement can be measured in various ways with various recording techniques. Here we used MEG and focussed on the reinstatement of temporal patterns, as measured with phase in lower frequencies (<40 Hz). We chose this approach over other approaches (i.e. based on power distribution across sensors; Jafarpour, et al. (2014); Fuentemilla, Penny, Cashdollar, Bunzeck, & Düzel (2010)) because previous studies showed that phase based temporal similarity measures effectively captures the reactivation of temporal patterns in memory (S. Michelmann, et al., 2016; Sebastian Michelmann, Bowman, et al., 2018; Sebastian Michelmann, Staresina, et al., 2018; Schreiner, et al., 2018; Staudigl, et al., 2015). We hypothesize to find higher temporal pattern similarity for successfully remembered items compared to forgotten items when the encoding and retrieval modalities match, and a reversal of this pattern when the encoding and retrieval modalities do not match. To foreshadow our findings, on a behavioural level we replicated the modality match effect (Mulligan and Osborn, 2009) for aurally encoded words, but not for visually encoded words (Figure 3). Therefore, we analysed the MEG data separately based on encoding modality and expected the above described interaction pattern only for the condition which showed a behavioural effect (i.e. for aurally encoded words, Figure 1B).

Methods

Ethics statement

The study was approved by the Institutional Review Board of the University of Konstanz. All participants gave written informed consent before the start of experiment in accordance with the Declaration of Helsinki.

Participants

24 participants (between 19 and 26 years old; mean age = 22 years; 17 female; 21 right-handed) were recruited for study. After excluding four participants (technical problems, excessive environmental noise), data from 20 participants are presented here. Participants received course

credits or monetary compensation for participation. All participants were German native speakers and reported no history of neurological disease and normal or corrected-to-normal vision.

Parts of this data have been published in (Staudigl and Hanslmayr, 2013) and (Westner, Dalal, Hanslmayr, & Staudigl, 2018), with respect to independent research questions and analyses (Staudigl & Hanslmayr, 2013: subsequent memory effects, study phase data only; Westner, Dalal, Hanslmayr & Staudigl, 2018: classification of stimulus modality, study phase data only).

Procedure

The experiment consisted of a study and a test phase. An outline of the procedure is depicted in Figure 1A. Before the start of the experiment, participants were instructed to count the syllables of each presented word and indicate via button press whether the word had two or more / less than two syllables. Participants were not instructed about the subsequent memory test, i.e., incidental encoding can be assumed. A short practice run ensured that the participants understood the task requirements.

During the study phase, words were presented in one of two modalities. In the visual encoding condition, words were projected on a screen. In the auditory encoding condition, vocal recordings of the words were presented via nonferromagnetic tubes to both ears, whilst the fixation cross remained on the screen. The duration of the word presentation was determined by the individual duration of the respective audio file, i.e., the time to pronounce the word (i.e. “dog” was presented for a shorter duration than “table”; mean duration = 697 ms, s.d. = 119 ms). Words were followed by a fixation cross. The duration of the word and fixation cross added up to 2000 ms. Thereafter, a question mark prompted the subject’s response in the syllable counting task. Participants were instructed before the start of the experiment to count the syllables of the word and indicate via button press whether the word had two syllables. The presentation of the question mark was either ended by the subject's button press or after a maximum duration of 1500 ms. A fixation cross (variable duration between 1000 and 1500 ms) preceded each word. After the study phase, participants performed a distracter task during which they counted backwards (in steps of three from a three-digit number) for 45 seconds.

After the distracter task, participants were informed about the upcoming, surprise recognition memory test. Participants were instructed to indicate their confidence on whether the item was old (presented during the study phase) or new (not presented during the study phase) on a six-point scale ranging from 'very sure old' to 'very sure new'. A short practice run ensured that the participants understood the task requirements.

During the test phase, words were presented with individual duration (determined by the duration of the respective audio file), followed by a fixation cross (summed duration of word and fixation cross = 2000 ms), followed by a stimulus picture depicting the response option. This stimulus prompted the participants' response indicating their confidence on whether they recognized the word as old or new, by pressing one out of six specified buttons on the response panel. The stimulus picture prompting the response was shown until the participants gave a response. A fixation cross (variable duration between 750 and 1250 ms) preceded each word.

Design and materials

420 unrelated German nouns were grouped into three lists with 140 words, each. Half of each list's words had two syllables, the other half had one, three or four syllables. Two lists were presented during the study phase (280 old items). Half of these words were presented visually,

the other half aurally. The assignment of words to modality was counterbalanced across participants. At test, the 280 old items intermixed with the 140 words from the remaining list (new items) were presented. Half of the old words were presented in the same modality (match condition), the other half in the different modality (mismatch condition) as during the study phase. Half of the new items were presented visually, the other half aurally.

The assignment of the lists to either study or test phase was counterbalanced across participants. For half of the participants, the confidence judgment during recognition test ranged from ‘very sure old’ (1) to ‘very sure new’ (6), for the other half from ‘very sure new’ (1) to ‘very sure old’ (6), thereby counterbalancing the response lateralization across subjects. Items were presented in a random order, with the constraint that not more than five words of the same modality (visual, auditory) and not more than five words from the same condition (match, mismatch) were presented sequentially.

The assignment of mismatch trials to either the auditory or visual mismatch condition (see Fig. 1b) was based on the modality of the study phase (i.e., mismatch trials were not averaged across conditions). Presenting a word aurally versus visually in the study phase results in modality-specific encoding patterns of neuronal activity. To investigate how the reactivation of these modality-specific neuronal patterns would affect retrieval during the test phase, it is important to keep the encoding modality constant and vary the modality of the retrieval cue (match or mismatch). Moreover, by keeping separate mismatch conditions for each we retained balanced trial numbers, thus keeping the signal-to-noise ratio for match and mismatch conditions constant.

Simulation

A simple computational model was generated to illustrate the opposing effects that the reactivation of neural patterns may have on memory performance in match and mismatch conditions. The model was inspired by a theoretical paper by E. Tulving (Tulving, 1983) who envisioned an early resonance process, termed “Ecphory” which can be described mathematically via a correlation between a stored engram and a retrieval cue and which inputs into the retrieval process. The simulation was programmed in MATLAB (version R2016a) and is a slightly modified version of a previous simulation published in (Staudigl, et al., 2015). The code for the simulation is available on https://github.com/hanslmayr/Audivis_Simulation. The model is divided into two layers, a semantic layer and a sensory layer, to simulate item meaning and the sensory modality in which an item is presented. Therefore, the sensory layer is further divided into auditory and visual units (see Figure 2). The semantic layer has 12 units; the sensory layer had 6 units for visual and 6 units for auditory patterns. Item patterns were expressed by a combination of 3 activated units on each layer (see Figure 2); 100 patterns were generated to simulate encoding and retrieval of 100 items. Thereby, each item was represented by a unique combination of a 3 unit pattern in a semantic layer and a 3 unit pattern in a sensory layer. These patterns were expressed in forms of 3 by 4 matrices, containing zeros for non-activated units and ones for activated units (see Figure 2).

Memory retrieval was simulated via an interaction, i.e. correlation, between a *reactivated pattern (R)* and a *cued pattern (C)* (Tulving, 1983). The simulation was calculated individually for each item, resulting in 100 trials. Within each trial, the reactivated pattern was first calculated separately for the semantic and the sensory layer. On the semantic layer, the reactivated semantic pattern (R_{sem}) was simply assumed to be the same as the cued pattern (C_{sem}) with a constant small amount of white noise (range 0 to 0.1). No difference was assumed between match and mismatch conditions, as in our experiment items were identical on a semantic level between match and mismatch conditions. On the sensory layer, we calculated the reactivated pattern (R_{sens}) by taking the original sensory pattern and adding white noise onto it. The strength of noise was controlled by a reactivation parameter (high vs. low reactivation) in order to simulate the effects of strength of reactivation on memory performance (i.e. high levels of noise result in low levels of reactivation). The cued pattern (C_{sens}) on the sensory layer varied between match and mismatch conditions. For the match condition, C_{sens} was the original pattern in the respective sensory layer in which the item has been encoded (i.e. auditory layer for aurally encoded words). Therefore, for match conditions C_{sens} overlapped with R_{sens}, with the strength of this overlap depending on the noise parameter which controls the strength of reactivation. For mismatch conditions, C_{sens} was taken from different units of the sensory layer (i.e. visual units for aurally encoded words). Therefore, for mismatch conditions C_{sens} never overlapped with R_{sens}. The overlap for a given trial was operationalized as the average across the fisher-z transformed 2d correlations (rz) between the cued and reactivated patterns on the semantic and sensory layers as per the below equation. Memory (Mem) performance is assumed to be a readout of the overlap:

$$Mem = \frac{rz(C_{sem}, R_{sem}) + rz(C_{sens}, R_{sens})}{2}$$

Data acquisition and preprocessing

MEG data was recorded with a whole-brain 148-channel magnetometer system (MAGNESTM 2500 WH, 4D Neuroimaging, San Diego, USA) inside a magnetically shielded room from participants in supine position. The data was continuously recorded at a sampling rate of 678.17

Hz and bandwidth of 0:1-200 Hz. Preprocessing of the data was done using the fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011), an open-source MATLAB toolbox for MEEG data analysis. Data from the study and test phase was epoched into single trials, with epochs ranging from 1500 ms prior to the onset of word presentation to 4000 ms after onset of word presentation. Trials were visually inspected for artefacts, to reject contaminated trials. Thereafter, independent component analysis (ICA) was used to correct for blinks, eye movements, and cardiac artefacts. On average, 3 (std = .79) components were rejected manually in the study phase, and 5.2 (std = 3.72) components were rejected in the test phase.

After artefact rejection, 54.95 (std = 6.1) trials remained on average in the auditory match condition (remembered trials: mean = 38.65, std= 8.2; forgotten trials: mean = 16.3 std= 6.59) and 55.4 (std = 6.15) in the auditory mismatch condition (remembered trials: mean = 35.1, std= 7.59; forgotten trials: mean = 20.3 std= 4.75). In the visual match condition, 55.3 (std= 6.3) trials remained (remembered trials: mean = 35, std= 8.06; forgotten trials: mean = 20.3 std= 6.37) and 53.4 (std = 6.7) in the visual mismatch condition (remembered trials: mean = 32.85, std= 7.35; forgotten trials: mean = 20.5 std= 5.66).

All trials remaining after artefact inspection were categorized according to the behavioural performance of each participant's during the recognition test phase. Trials including old items that were confidently judged as old (responses 1, 2, and 3) constituted remembered trials, the remaining trials including old items were classified as forgotten trials, in both the study and the test phase. Trials including new items that were confidently judged as being new (responses 4, 5, and 6) constituted correct rejections, the remaining trials including new items were classified as false alarms.

Similarity analysis

In order to identify reinstatement of neuronal patterns from encoding during retrieval, similarity in neuromagnetic activity between words presented at study and test was assessed. Based on previous findings (S. Michelmann, et al., 2016; Staudigl, et al., 2015), we focused on similarity in oscillatory phase, using the pairwise phase consistency index (PPC; (Vinck, van Wingerden, Womelsdorf, Fries, & Pennartz, 2010). PPC estimates the phase consistency between two separate signals across trials, quantifying the extent of a consistent phase relationship among them. Compared to other measures of phase consistency (e.g., phase-locking value), PPC is advantageous since it is not biased by the number of trials. As a consequence of the unbiasedness, the PPC can, however, have negative values (Vinck, et al., 2010).

In order to provide the necessary phase information for the computation of the PPC, time-frequency representations of the data were computed by a sliding time window approach with a window length of 0.5 s in steps of 50 ms across the data. After multiplying a hanning taper to each window, the Fourier transformation was calculated for frequencies between 4 and 40 Hz in steps of 2 Hz.

The PPC was calculated between each of the time-frequency bins in the study phase and the corresponding time-frequency bins in the test phase, separately for remembered and forgotten trials, match and mismatch condition, and auditory and visual encoding condition.

Statistics

Statistical quantification of the data was performed by a cluster-based nonparametric permutation approach (Maris and Oostenveld, 2007) identifying clusters of activity on the basis of rejecting the null hypothesis while controlling for multiple comparisons over sensors, time-points and frequencies. For each time and frequency bin at each sensor, a test statistic was calculated (10,000 permutations), based on a paired samples t-test comparing the difference in PCC in the match condition (remembered minus forgotten trials) to the difference in PPC in the mismatch condition (remembered minus forgotten trials; see Figure 1b).

T-values above the cluster-forming threshold ($p < 0.05$, two-sided t-test) were clustered based on adjacency in time, frequency and space (a minimum of 2 adjacent sensors was required for forming a cluster). T-statistics were summed in each cluster and compared against the distribution of maximal clusters provided by the permutation approach, using the Monte Carlo method for randomly assigning the differences (remembered minus forgotten) to conditions (match, mismatch). Only the cluster with the largest summed value was considered and tested against the permutation distribution. The null-hypothesis that the match and mismatch condition showed no difference in PPC was rejected at an alpha-level of 0.05 (two-tailed).

Results

Behavioural results

A summary of the behavioural results is depicted in Figure 3. Averaged across modality conditions, participants correctly recognized significantly more words ($t_{19} = 3.56$, $p < 0.005$) in the match (mean = .661, SE = .023) than the mismatch condition (mean = .621, SE = .020). The interaction between modality (auditory vs. visual) and test condition (match vs. mismatch) was not significant ($F_{1,19} = 2.72$, $p > .115$).

In the auditory encoding condition, participants correctly recognized significantly more words in the match (mean = .695, SE = .025) than in the mismatch (mean = .63, SE = .023) condition (t_{19} = 3.55, p < .005). No significant difference between correctly recognized words in the match (mean =.626, SE = .024) and the mismatch (mean = .612, SE = .024) condition was found for visually encoded items (t_{19} = .73, p > 0.475). Note that this pattern of results cannot be interpreted as a modality x condition interaction effect (see Nieuwenhuis, Forstmann, & Wagenmakers, 2011, for an important discussion on interaction effects).

The false alarm rate was not significantly different (t_{19} = -0.62, p > .54) for items presented visually (mean = .26, SE = .027) or aurally (mean = .27, SE = .025) during test.

When only including the high confidence judgements (“very sure old”), the interaction between modality (auditory vs. visual) and test condition (match vs. mismatch) was significant ($F_{1,19}$ = 10.79, p < .005). In the auditory encoding condition, participants correctly recognized significantly more words in the match (mean = .384, SE = .038) than in the mismatch (mean = .307, SE = .035) condition (t_{19} = 3.70, p < .005). No significant difference between correctly recognized words in the match (mean =.285, SE = .032) and the mismatch (mean = .311, SE = .032) condition was found for visually encoded items (t_{19} = -1.43, p > 0.168).

Simulation results

We created a simple computational model to formalize our hypotheses and to illustrate the impact of neural pattern reactivation on memory performance for match and mismatch conditions. The results of this model are shown in Figure 2. If the modality of the retrieval cue matches the encoding modality (Figure 2a), high levels of sensory reactivation lead to a strong resonance, i.e. a better memory performance. On the other hand, if the modality of the retrieval cue does not match the modality of the encoded pattern, then high levels of sensory reactivation lead to lower resonance, i.e. worse memory performance (Figure 2b). This latter effect arises because in the mismatch modality the patterns of the engram and the cued patterns never overlap as they are represented in different units. Reactivation of a sensory trace in the mismatch condition therefore leads to a negative correlation which then reduces the overall correlation between the cue and the reactivated pattern. This simple model illustrates what we should expect to find if neural pattern reactivation underlies the encoding specificity principle. That is, high levels of reactivation for remembered trials in the match condition, but low levels of reactivation for remembered trials in the mismatch condition.

Reinstatement of auditory temporal patterns interacts with modality match

In order to identify memory reinstatement of temporal patterns from encoding during retrieval, similarity in neuromagnetic activity between words presented at study and test was assessed using the pairwise phase similarity (S. Michelmann, et al., 2016). Given the asymmetric behavioural pattern, i.e. a modality match effect was only observed for aurally encoded words, the analysis was carried out separately for visually and aurally encoded words.

In the auditory encoding condition, cluster-based nonparametric statistics yielded a significant interaction effect (p cluster $< .05$), indicating that the difference in the match condition (remembered – forgotten) was reliably higher than in the mismatch condition (remembered – forgotten) at 6-8 Hz between 0.15 and .2 seconds after word onset (Figure 4a). Figure 4b depicts a topography of the interaction effect, highlighting left central sensors contributing to the significant cluster.

Post-hoc t-tests revealed that the neuronal similarity was higher for the remembered than the forgotten words in the match condition for ($t_{19} = 5.02$, $p < .0001$), whereas in the mismatch condition, similarity was higher for the forgotten than the remembered words ($t_{19} = 2.11$, $p < .05$, see Figure 4b). Comparing only remembered words revealed that the neuronal similarity was higher in the match than in the mismatch condition ($t_{19} = 2.46$, $p < .05$). It should be noted that this post-hoc analysis is statistically biased by selectively averaging across those channels, time and frequency bins that show a significant interaction. Nevertheless, this post-hoc analysis adds information as it shows that the interaction is not just driven by a difference in one condition, but by opposing differences between remembered and forgotten items in both conditions.

In another post-hoc analysis, we included high confidence trials (“sure old” responses) in the remembered condition only (HCrem). When averaging across the time and frequency bins that showed a significant interaction in the previous analysis, a very similar topography was found (Fig 5, left). When averaging across the channels, time and frequency bins that showed a significant interaction in the previous analysis, the interaction effect for the HCrem items resembled the pattern found in the original analysis: The interaction (match/mismatch vs. remembered/forgotten) was significant ($F_{1,19}=9.76$, $p < 0.01$), as well as the comparison between HCrem and forgotten items in the match condition ($t_{19} = 3.72$, $p < .005$). The comparison between HCrem and forgotten items in the mismatch condition was not significant ($t_{19} = -1.13$, $p > .27$), as well as the comparison between HCrem in the match and the mismatch condition ($t_{19} = 1.39$, $p > .18$).

In the visual condition, there were no significant interaction effects when comparing the difference (remembered – forgotten) in the match condition to the difference (remembered – forgotten) in the mismatch condition (all cluster p's > .59).

When remembered versus forgotten items were compared irrespective of match / mismatch condition, no significant effects were found in the auditory (p cluster > .08) nor in the visual condition (p cluster > .16).

Discussion

In the current study we investigated how the reactivation of encoding patterns interacts with memory in the face of retrieval cues which match or do not match the encoding modality. For aurally encoded words, we show that reactivation of encoding patterns is related to opposing memory outcomes; depending on whether the retrieval cue matches the encoding modality or not. Specifically, reactivation of a sensory encoding pattern is beneficial for memory when the retrieval cue matches the encoding modality, but is detrimental when the cue does not match the encoding modality. Together with our previous findings (Staudigl, et al., 2015), these results provide insights into the neural basis underlying a classic effect in cognitive psychology, i.e. Encoding Specificity or Transfer Appropriate Processing (Morris, et al., 1977; Tulving and Thomson, 1973): the more the retrieval cue overlaps with the encoded pattern, i.e. the stronger the two resonate, the more likely it is that the item will be recognized as old (Rugg, et al., 2008). Paradoxically, when the retrieval cue is presented in a different sensory modality than in which the item has been encoded, reactivation of sensory patterns was related to decreased memory performance. This pattern is in line with the outcome of a simple model which illustrated our results on a neural and behavioural level: the mismatch between the reactivated sensory pattern (i.e. “dog” in auditory layer) and the sensory pattern provided by the retrieval cue (i.e. “dog” visually presented) decreases the overlap (or resonance) between the cue and the reactivated pattern, and, thereby, the memory performance. This decrease is even more pronounced in case of a strongly reactivated memory trace.

Neural evidence consistent with Encoding Specificity comes from fMRI (Park and Rugg, 2008) and ERP studies (Bauch and Otten, 2012). These studies show that the neural correlates of successful encoding vary depending on whether the memory is later tested with cues that match or do not match the encoding modality (Park and Rugg, 2008). We here add to this neural evidence by using a temporal pattern similarity approach and show that reactivation of neural patterns established at encoding has opposing effects on memory depending on how the memory is cued. One advantage of the here used approach is that it returns a time-resolved measure of memory reactivation, which is allows to infer whether reactivation is an early process leading up to a memory judgement or whether it is a later by-product of a memory judgement (i.e. imagery, or retrieval monitoring). The early time window of reactivation observed here (i.e. 150 – 200 ms) supports the former view and suggests an early resonance process between a memory cue and a stored engram which occurs before the memory decision is being made, which has been termed Ecphory (Tulving, 1983). This finding is consistent with a larger body of studies describing similar early reactivation effects (Jafarpour, et al., 2014; Waldhauser, Braun, & Hanslmayr, 2016; Wimber, Maass, Staudigl, Richardson-Klavehn, & Hanslmayr, 2012, but see Lewis, Schriefers, Bastiaansen, & Schoffelen, 2018; Price and Johnson, 2018).

Interestingly, on a behavioural level the observed modality match effects were asymmetric, i.e. a modality match effect was observed only for aurally encoded words but not for visually encoded words (albeit it should be acknowledged that there was no significant 2-way interaction between match/mismatch and modality). Consistent with this asymmetric pattern, only aurally encoded items showed an interaction effect in terms of neural pattern reactivation with modality match, whereas no such effect was obtained for visually encoded items. A possible post-hoc explanation for this asymmetry is that the visual presentation of words automatically evokes auditory patterns encoded by the participants. This would be consistent with reports of high levels of auditory cortex activation typically observed during reading, and in line with the idea that auditory cortex is part of the brain's "reading network" (Wandell and Le, 2017). This activation of auditory patterns might have been further promoted by the here used encoding strategy, i.e. syllable counting which arguably enforces subjects to covertly pronounce the word. In such a scenario, for a visually encoded word, an auditory cue is just as effective as a visual cue because the engram contains both patterns. Consistent with this explanation, a previous study (Mulligan and Osborn, 2009) which reported modality match effects for visually *and* aurally encoded words used a different encoding strategy, i.e. intentional encoding (as opposed to syllable counting and incidental encoding employed in our study). On a more general level, these considerations highlight the importance of the cognitive processes carried out at encoding and retrieval which determine memory effects observed on a behavioural and neural level (Hanslmayr and Staudigl, 2014; Rugg, et al., 2008).

The here observed reactivation of auditory patterns occurred in a frequency range of 6-8 Hz. This frequency range perfectly matches a previous study where we found that dynamic auditory patterns are encoded in the phase of a 6-8 Hz oscillation (S. Michelmann, et al., 2016; Sebastian

Michelmann, Bowman, et al., 2018; Schreiner, et al., 2018). This result is also in line with another study showing that auditory stimuli can be decoded from neural phase patterns at 4-8 Hz (Ng, Logothetis, & Kayser, 2013). This result is, however, in contrast to a previous study (Staudigl, et al., 2015) investigating match-mismatch with background movies, where we found reactivated patterns at a higher frequency range (i.e. 30 Hz). This difference could be either due to the different sensory modalities of the stimuli between the two studies (i.e. auditory vs visual), or could be due to slightly different analysis approaches. In contrast to Staudigl et al. 2015, we here used pairwise phase consistency (PPC) as we did in Michelmann et al (2016, 2018) which has the advantage of resulting in a time-resolved measure of Temporal Pattern Similarity. Together with previous studies (S. Michelmann, et al., 2016; Sebastian Michelmann, Bowman, et al., 2018; Ng, et al., 2013; Schreiner, et al., 2018) these results suggest a general role of 6-8 Hz oscillations for coding auditory information not only on a perceptual level, but also in memory.

Conclusion

The encoding specificity principle or transfer appropriate processing are classic frameworks which have influenced memory research for decades. In their seminal paper published in 1973, Tulving and Thomson describe the basic ideas of Encoding Specificity but also state:

„The terms are ill defined, and the concepts do not explain too much at this time. Yet they serve to remind us that something else besides the properties of a presented item determines how well the item is remembered and that an important research problem is to find what this something else is and how it works“ (Tulving and Thomson, 1973).

More than 4 decades later we still know little about the mechanisms underlying the encoding specificity principle. In neural terms, the encoding specificity principle can be expressed as a resonance process between a cue and a stored engram, whereby the overlap between the two determines whether a memory can be retrieved or not (Rugg, et al., 2008; Tulving, 1983). We here show that the strength of reactivation of that memory trace plays a central role in this process and thus is a critical ingredient underlying Encoding Specificity (or Transfer Appropriate Processing). We hope that future studies pick up on these ideas using multivariate analysis tools which allow to quantify the reactivation of stored memory traces (i.e. the engrams) using standard neuroimaging tools such as EEG, MEG or fMRI.

Figures & Figure Captions

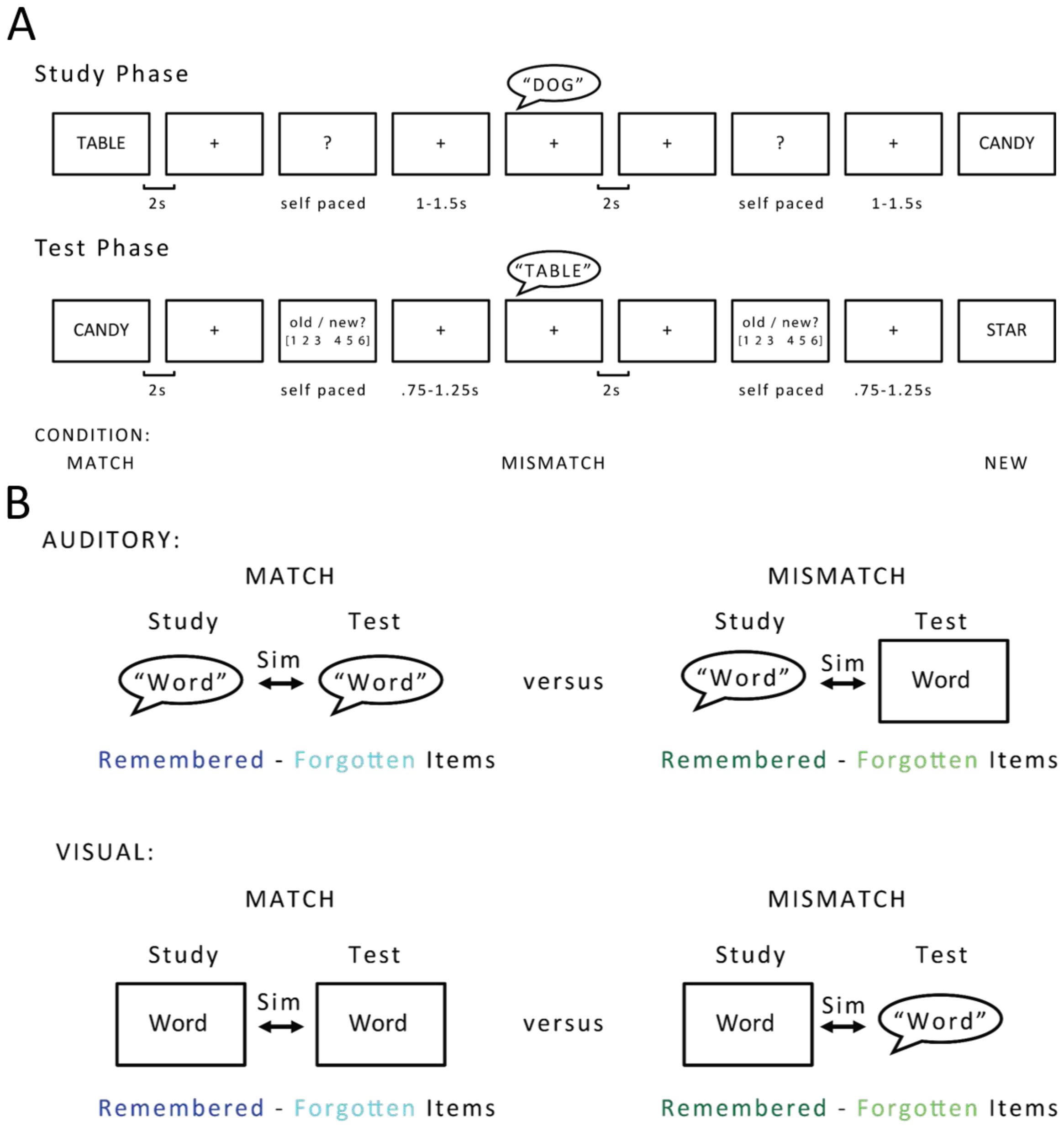

Figure 1.

Experimental Procedure and analysis rationale. A) Procedure. In the study phase, words were either presented visually or aurally (indicated by the speech bubble). Duration of presentation was adjusted individually for each word to match the duration of the audio file. The combined presentation time of the word and fixation cross was 2 always seconds. Participants were instructed to count the syllables (two or more/less than two) of each word and respond during the question mark. In the test phase, old words were either presented in the same (match condition) or different (mismatch condition) modality as during the study phase. Old words were randomly intermixed with new words, and participants were asked to judge their confidence as to whether the word was old or new on a six-point scale, ranging from ‘very sure old’ (1) to ‘very sure new’ (6). B) Analysis rationale. Temporal Pattern Similarity was computed for word pairs presented during study and test. Similarity differences between remembered and forgotten words were compared between the match and mismatch condition, separately for the auditory and the visual condition.

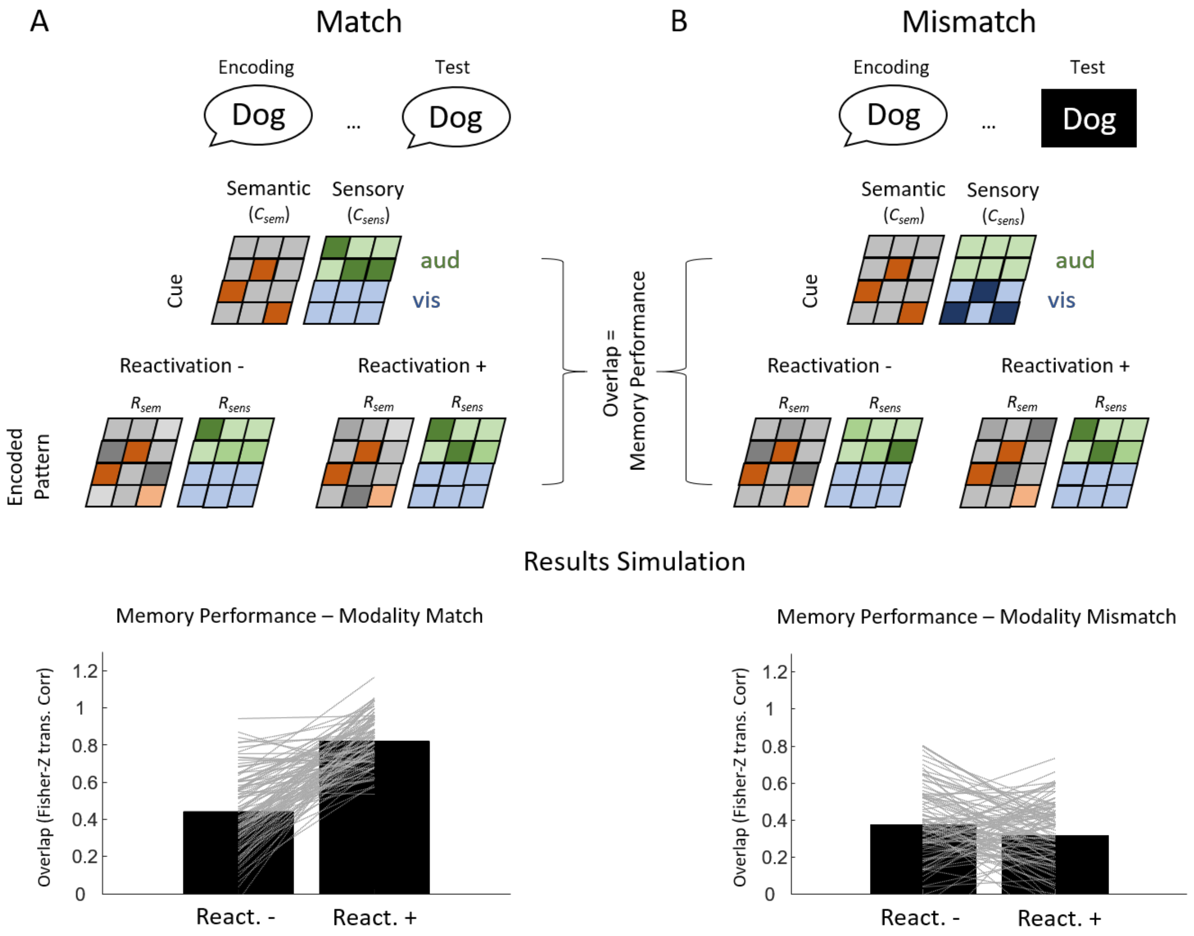

Figure 2. The architecture of a simple computational model (top panel) and results of the simulation (bottom panel) are shown for Match (A) and Mismatch (B) conditions. Individual patterns were generated on a semantic and a sensory layer to represent cued and encoded item patterns. The sensory layer was divided into visual and auditory units. Memory performance was conceived of as the overlap between a cued pattern (C_{sem}, C_{sens}) and a reactivated pattern (R_{sem}, R_{sens}) as calculated with 2d correlations. These correlations were averaged across the two layers. The bottom panel shows the overlap averaged across simulation runs (bar plots) and individual trials (grey lines) for high and low reactivation levels (React.+ and React.-, respectively). 100 simulation runs per condition and reactivation level were carried out.

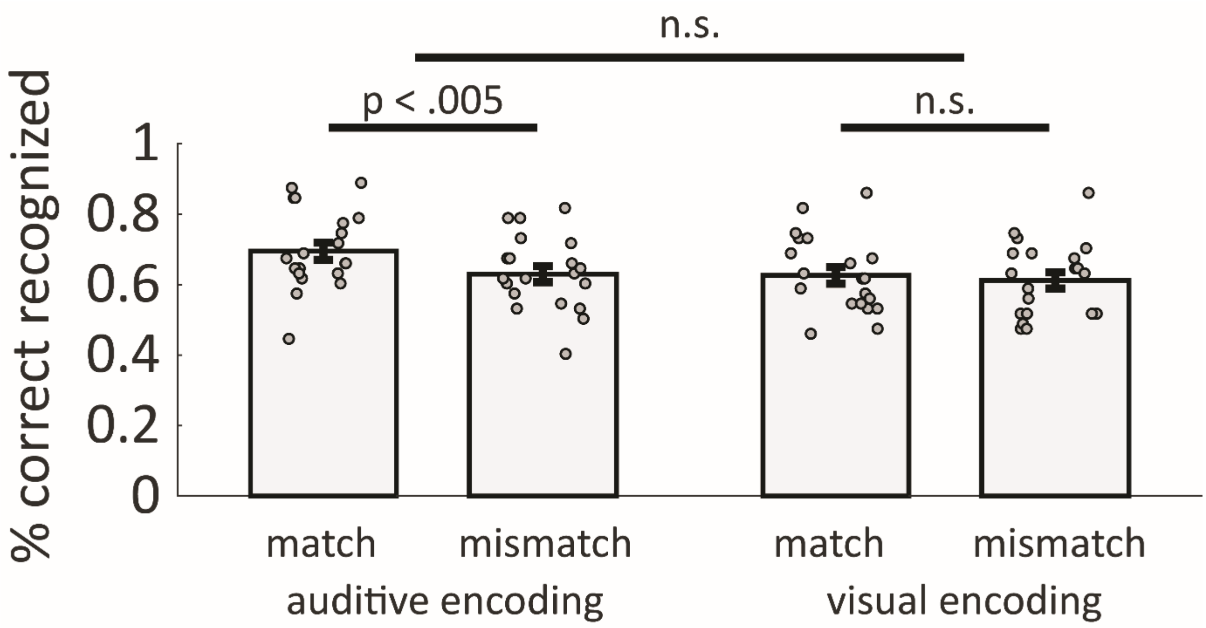

Figure 3. Behavioural results. Recognition performance in the match and mismatch condition as a function of auditory (left) and visual (right) encoding. A modality match effect was observed only for the auditory encoding condition ($p < .005$). Dots represent individual data points, errors bars depict S.E.

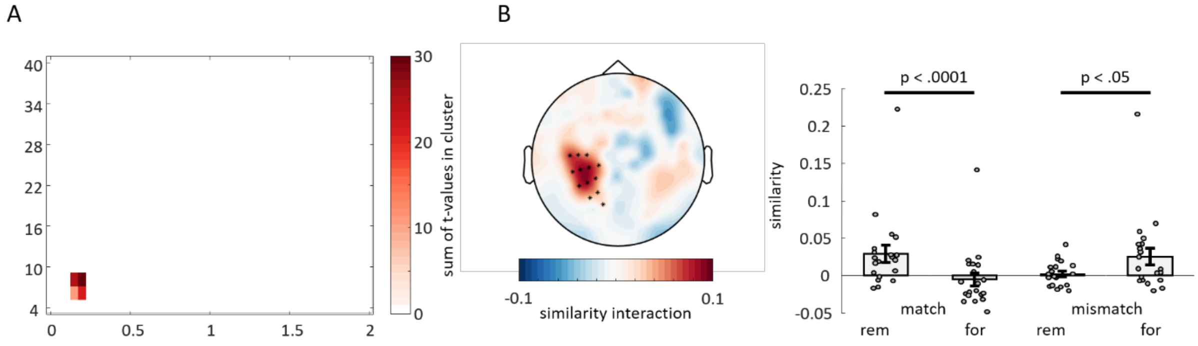

Figure 4. Interaction effect in the auditory condition. A) Significant interaction (p cluster < .05) between the match (remembered – forgotten) and the mismatch (remembered – forgotten) condition, at 6-8 Hz and .15-.2 s after word onset. B) Left: Topography of the interaction effect ([PPC match remembered – PPC match forgotten] - [PPC mismatch remembered – PPC

mismatch forgotten]), averaged across time and frequency depicted in A. Sensors contributing to the significant interaction are highlighted. Right: Similarity (PPC) for the remembered and forgotten words, in the match and mismatch condition, respectively, as averaged across time (.15-.2 s), frequencies (6-8 Hz) and sensors highlighted in the topography. Dots represent individual data points, errors bars depict S.E.

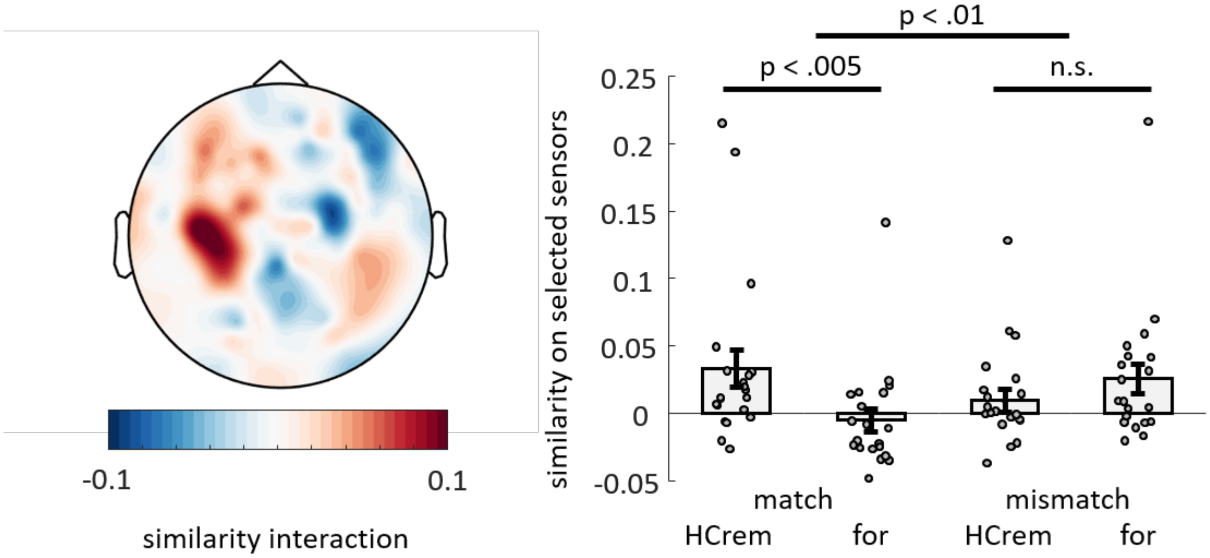

Figure 5. High confidence – remembered words (HCrem) in the auditory condition. A) Left: Topography of the interaction ([PPC match HCrem – PPC match forgotten] - [PPC mismatch HCrem – PPC mismatch forgotten]), averaged across time and frequency depicted in Figure 4A. Right: Similarity (PPC) for HCrem and forgotten words, in the match and mismatch condition, respectively, as averaged across time (.15-.2 s), frequencies (6-8 Hz) and sensors highlighted in the topography of Figure 4b. Dots represent individual data points, errors bars depict S.E.

Acknowledgements

We thank Ann-Kristin Rombach, Leona Hellwig, Marina Koepfer, and Janine Weichert for help with data acquisition. This research was supported by European Union's Horizon 2020, (https://ec.europa.eu/programmes/horizon2020/, 661373) to TS; Deutsche Forschungsgemeinschaft (http://www.dfg.de/, Emmy Noether Programme Grant HA 5622/1-1) to SH; European Research Council (https://erc.europa.eu/, Consolidator Grant 647954) to SH; Wolfson Foundation and Royal Society (https://royalsociety.org/grants-schemes-awards/grants/wolfson-research-merit/) to SH.

References

Bauch, E. M., & Otten, L. J. (2012). Study-test congruency affects encoding-related brain activity for some but not all stimulus materials. *J Cogn Neurosci, 24*(1), pp. 183-195. doi:10.1162/jocn_a_00070 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21671740

Fuentemilla, L., Penny, W. D., Cashdollar, N., Bunzeck, N., & Düzel, E. (2010). Theta-coupled periodic replay in working memory. *Current Biology, 20*(7), pp. 606-612.

Godden, D. R., & Baddeley, A. D. (1975). Context-Dependent Memory in 2 Natural Environments - Land and Underwater. *British Journal of Psychology, 66*(Aug), pp. 325-331. doi:DOI 10.1111/j.2044-8295.1975.tb01468.x Retrieved from <Go to ISI>://WOS:A1975AM81800007

Hanslmayr, S., & Staudigl, T. (2014). How brain oscillations form memories--a processing based perspective on oscillatory subsequent memory effects. *Neuroimage, 85 Pt 2*, pp. 648-655. doi:10.1016/j.neuroimage.2013.05.121 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23769913

Jafarpour, A., Fuentemilla, L., Horner, A. J., Penny, W., & Duzel, E. (2014). Replay of very early encoding representations during recollection. *J Neurosci, 34*(1), pp. 242-248. doi:10.1523/JNEUROSCI.1865-13.2014 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24381285

Johnson, J. D., McDuff, S. G., Rugg, M. D., & Norman, K. A. (2009). Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron, 63*(5), pp. 697-708. doi:10.1016/j.neuron.2009.08.011 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19755111

Kerrén, C., Linde-Domingo, J., Hanslmayr, S., & Wimber, M. (2018). An optimal oscillatory phase for pattern reactivation during memory retrieval. *Current Biology, 28*(21), pp. 3383-3392. e3386.

Lewis, A. G., Schriefers, H., Bastiaansen, M., & Schoffelen, J.-M. (2018). Assessing the utility of frequency tagging for tracking memory-based reactivation of word representations. *Scientific reports, 8*

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods, 164*(1), pp. 177-190. doi:10.1016/j.jneumeth.2007.03.024 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17517438

Mcdermott, K. B., & Roediger, H. L. (1994). Effects of Imagery on Perceptual Implicit Memory Tests. *Journal of Experimental Psychology-Learning Memory and Cognition, 20*(6), pp. 1379-1390. doi:Doi 10.1037/0278-7393.20.6.1379 Retrieved from <Go to ISI>://WOS:A1994PP52100009

Michelmann, S., Bowman, H., & Hanslmayr, S. (2016). The Temporal Signature of Memories: Identification of a General Mechanism for Dynamic Memory Replay in Humans. *PLoS Biol, 14*(8), p e1002528. doi:10.1371/journal.pbio.1002528 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27494601

Michelmann, S., Bowman, H., & Hanslmayr, S. (2018). Replay of stimulus-specific temporal patterns during associative memory formation. *Journal of cognitive neuroscience, 30*(11), pp. 1577-1589.

Michelmann, S., Staresina, B. P., Bowman, H., & Hanslmayr, S. (2018). Speed of time-compressed forward replay flexibly changes in human episodic memory. *Nature Human Behaviour*doi:10.1038/s41562-018-0491-4 Retrieved from https://doi.org/10.1038/s41562-018-0491-4

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of Processing Versus Transfer Appropriate Processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), pp. 519-533. doi:Doi 10.1016/S0022-5371(77)80016-9 Retrieved from <Go to ISI>://WOS:A1977EA05100001

Mulligan, N. W., & Osborn, K. (2009). The modality-match effect in recognition memory. *J Exp Psychol Learn Mem Cogn, 35*(2), pp. 564-571. doi:10.1037/a0014524 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19271869

Ng, B. S., Logothetis, N. K., & Kayser, C. (2013). EEG phase patterns reflect the selectivity of neural firing. *Cereb Cortex, 23*(2), pp. 389-398. doi:10.1093/cercor/bhs031 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22345353

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. *Nature neuroscience, 14*(9), p 1105.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci, 2011*, p 156869. doi:10.1155/2011/156869 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21253357

Park, H., & Rugg, M. D. (2008). The relationship between study processing and the effects of cue congruency at retrieval: fMRI support for transfer appropriate processing. *Cereb Cortex, 18*(4), pp. 868-875. doi:10.1093/cercor/bhm130 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17652467

Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science, 310*(5756), pp. 1963-1966. doi:10.1126/science.1117645 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16373577

Price, M. H., & Johnson, J. D. (2018). Failure to reactivate salient episodic information during indirect and direct tests of memory retrieval. *Brain research, 1699*, pp. 9-18.

Rugg, M. D., Johnson, J. D., Park, H., & Uncapher, M. R. (2008). Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. *Prog Brain Res, 169*, pp. 339-352. doi:10.1016/S0079-6123(07)00021-0 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18394485

Schreiner, T., Doeller, C. F., Jensen, O., Rasch, B., & Staudigl, T. (2018). Theta Phase-Coordinated Memory Reactivation Reoccurs in a Slow-Oscillatory Rhythm during NREM Sleep. *Cell Reports, 25*(2), pp. 296-301.

Smith, S. M., & Manzano, I. (2010). Video context-dependent recall. *Behav Res Methods, 42*(1), pp. 292-301. doi:10.3758/BRM.42.1.292 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20160308

Staresina, B. P., Henson, R. N., Kriegeskorte, N., & Alink, A. (2012). Episodic reinstatement in the medial temporal lobe. *J Neurosci, 32*(50), pp. 18150-18156. doi:10.1523/JNEUROSCI.4156-12.2012 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23238729

Staresina, B. P., Michelmann, S., Bonnefond, M., Jensen, O., Axmacher, N., & Fell, J. (2016). Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *Elife, 5*doi:10.7554/eLife.17397 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27508355

Staudigl, T., & Hanslmayr, S. (2013). Theta oscillations at encoding mediate the context-dependent nature of human episodic memory. *Curr Biol, 23*(12), pp. 1101-1106. doi:10.1016/j.cub.2013.04.074 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23746635

Staudigl, T., Vollmar, C., Noachtar, S., & Hanslmayr, S. (2015). Temporal-pattern similarity analysis reveals the beneficial and detrimental effects of context reinstatement on human memory. *J Neurosci, 35*(13), pp. 5373-5384. doi:10.1523/JNEUROSCI.4198-14.2015 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25834061

Tulving, E. (1983). Ecphoric Processes in Episodic Memory. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 302*(1110), pp. 361-371. doi:DOI 10.1098/rstb.1983.0060 Retrieved from <Go to ISI>://WOS:A1983RE67600009

Tulving, E., & Thomson, D. M. (1973). Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review, 80*(5), pp. 352-373. doi:DOI 10.1037/h0020071 Retrieved from <Go to ISI>://WOS:A1973Q939300003

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., & Pennartz, C. M. (2010). The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. *Neuroimage, 51*(1), pp. 112-122. doi:10.1016/j.neuroimage.2010.01.073 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20114076

Waldhauser, G. T., Braun, V., & Hanslmayr, S. (2016). Episodic Memory Retrieval Functionally Relies on Very Rapid Reactivation of Sensory Information. *J Neurosci, 36*(1), pp. 251-260. doi:10.1523/JNEUROSCI.2101-15.2016 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26740665

Wandell, B. A., & Le, R. K. (2017). Diagnosing the Neural Circuitry of Reading. *Neuron, 96*(2), pp. 298-311. doi:10.1016/j.neuron.2017.08.007 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/29024656

Westner, B. U., Dalal, S. S., Hanslmayr, S., & Staudigl, T. (2018). Across-subjects classication of stimulus modality from human MEG high frequency activity. *PLoS Computational Biology, in press*

Wimber, M., Maass, A., Staudigl, T., Richardson-Klavehn, A., & Hanslmayr, S. (2012). Rapid memory reactivation revealed by oscillatory entrainment. *Curr Biol, 22*(16), pp. 1482-1486. doi:10.1016/j.cub.2012.05.054 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22795695

Yaffe, R. B., Kerr, M. S., Damera, S., Sarma, S. V., Inati, S. K., & Zaghloul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proc Natl Acad Sci U S A, 111*(52), pp. 18727-18732. doi:10.1073/pnas.1417017112 Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25512550