Research Report

Electrophysiological correlates of verbal and tonal working memory

Katrin Bittrich[a,], Katrin Schulze[b,c], Stefan Koelsch[b,d]*

[a]*Martin-Luther-University Halle-Wittenberg, Department of Psychology, 06099 Halle (Saale), Germany*
[b]*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany*
[c]*Developmental Cognitive Neuroscience Unit, UCL Institute of Child Health, London, UK*
[d]*Cluster of Excellence 'Languages of Emotion', Freie Universität Berlin, Berlin, Germany*

ARTICLE INFO

ABSTRACT

The present study's basic research question concerns differences in the processing of verbal and tonal stimuli in working memory. Participants had to rehearse sequences containing tonal and verbal information and to decide whether a subsequently presented probe stimulus belonged to the previously presented sequence. Electrophysiological measures were taken and analysed with regard to local (event-related potentials, ERP) and global (microstates) aspects. A larger N400 amplitude occurred for new compared to old items in the verbal, but not the tonal condition. The microstate analysis revealed differences in the specificity of several microstates for old compared to new items in the tonal and the verbal working memory condition. Corroborating previous results, the present study reveals differences in the efficiency of working memory processes for tonal compared to verbal stimuli with processes being more capable for the verbal compared to the tonal condition.

1. Introduction

A temporary storage system being in charge of information storage and processing, as well as decision making and problem solving, was proposed by Atkinson and Shiffrin (1968). Besides a storage component (short term store, STS), the model comprises control processes, for example rehearsal, coding, decision making, and retrieval strategies (Atkinson and Shiffrin, 1971). Retrieval processes from STS are rather fast as confirmed by memory scanning experiments by Sternberg (1966; c.f. Atkinson and Shiffrin, 1971). A more comprehensive model implying all the Atkinson and Shiffrin STS mechanisms, but extending the model to a multicomponent store,

is the working memory model (WM) proposed by Baddeley and Hitch (1974; see also Baddeley, 1992, 2003). It comprises an attentional control system (the central executive) which works together with the three sub-systems: the phonological loop, the visuo-spatial sketchpad, and the episodic buffer. The phonological loop, in turn, consists of two subsystems: a phonological store which holds auditory information for a short period of time, and an articulatory control process which serves as rehearsal mechanism (comparable with subvocal speech; Baddeley, 2003).

Researchers have asked the question whether the same system or different systems underlie the temporal storage of verbal and tonal information. However, behavioural studies

---

* *Corresponding author at:* Matrin-Luther-University Halle-Wittenberg, Department of Psychology, 06099 Halle (Saale), Germany. Fax: +49 345 27060.
E-mail address: katrin.bittrich@psych.uni-halle.de (K. Bittrich).

investigating the processing of verbal and tonal material (Chan et al., 1998; Deutsch, 1970; Lee, 2002; Mondor and Morin, 2004; Pechmann and Mohr, 1992 ; Salamé and Baddeley, 1989; Schendel and Palmer, 2007; Semal et al., 1996) have not yet provided a consistent answer to this question. Deutsch (1970) presented different types of intervening material (such as tones or phonemes) between a standard and a comparison tone. The study revealed reduced performance for tonal WM when presented with intervening tones compared to intervening verbal material or silence. Because interference is expected to be stronger within a specialised system rather than between two different systems, these results favour the assumption that different systems underlie tonal and verbal WM. However, studies that investigated aspects of WM for speech and musical sounds by focussing on pitch similarity between task relevant and interfering material (e.g., Semal et al., 1996), articulatory tonal and verbal suppression (e.g., Schendel and Palmer, 2007), or musical expertise (Pechmann and Mohr, 1992; Williamson et al., 2010) support the assumption of shared storage and processing resources underlying tonal and verbal WM.

Using functional magnetic resonance imaging (fMRI), three studies have directly compared the neural networks underlying tonal and verbal WM (Hickok et al., 2003; Koelsch et al., 2009; Schulze et al., 2011a, 2011b). These studies revealed that tonal WM substantially overlaps with the core structures of verbal WM, namely Broca's area, pre-motor cortex (Hickok et al., 2003; Koelsch et al., 2009; Schulze et al., 2011a, 2011b) and the planum temporale (Hickok et al., 2003; Koelsch et al., 2009). In addition, Schulze et al. (2011a, 2011b) also observed activation of different subcomponents during verbal and tonal WM in musicians compared to non-musicians. The combined results thus indicate that verbal and tonal WM largely overlap in non-musicians, and that two WM subsystems (a phonological and a tonal loop) serve the processing of verbal and tonal information in musicians.

The mentioned neuroimaging studies used fMRI to localise neural correlates of WM for the rehearsal mechanism. The present study aimed to investigate temporal aspects of memory scanning and decision making in tonal and verbal WM using EEG. According to our knowledge, this is the first study that compared electrophysiological correlates underlying auditory WM for verbal and musical information (for studies on WM processes for non-verbal environmental sounds see Chao et al., 1995; Chao and Knight, 1996; for a study on serial position effects using digits and musical notes see Patterson et al., 1991). Two prominent event-related potential (ERP) components have been observed in studies investigating neurophysiological aspects of WM, the P300 (Kok, 2001; Polich, 2007; Rugg and Coles, 1995) and the N400 (Curran, 2004; Kutas and Federmeier, 2011; Kutas and Hillyard, 1980).

Results from studies on the functional significance of the P300 in WM tasks are rather consistent: the P300 amplitude increases with increasing processing demands needed for the encoding and storage of information (for a review see Friedman and Johnson, 2000). During probe recognition, different aspects can affect the P300 amplitude, all of them more or less related to task difficulty. For example, amplitudes are smaller for positive probes with earlier serial position within the to-be-remembered sequence (Chao, and Knight, 1996; Golob and Starr, 2004; McEvoy et al., 1998; Patterson et al.,

1991) and amplitudes decline after long retention intervals, which is attributable to the decay of information in WM (Kim et al., 2001; Nielsen-Bohlman and Knight, 1994).

The functional significance of the N400 in the context of WM processes is rather heterogeneous. The N400 was originally investigated with regard to semantic-lexical processing: items that do not suit a previously established context elicit an N400 (Kutas and Federmeier, 2011; Kutas and Hillyard, 1980). Similar patterns have been described for recognition tasks with correctly identified old items eliciting smaller N400 amplitudes than correctly rejected new items (for a review see Friedman and Johnson, 2000; Kutas and Federmeier, 2011). Although the ERP old/new effect has first been reported for long-term memory tasks, it has been shown for short-term memory task as well (Crites et al., 1998, 2000). This ERP old/new effect might reflect context integration in such way that the positive probe item matches a previously established context (the memory set) whereas the negative probe item does not.

Following up the neuroimaging studies by Koelsch et al. (2009) and Schulze et al. (2011a, 2011b), the present EEG-study investigated the temporal dynamics of WM for auditory verbal and tonal material, using the same stimulus material as in the neuroimaging study by Schulze et al. (2011a, 2011b): verbal and tonal information was presented simultaneously in memory set sequences of five items, and participants were asked to rehearse either the verbal or the tonal information during a rehearsal period. Subsequently, participants had to decide whether a probe stimulus was an element of the memory set (positive probe) or not (negative probe, for details see Fig. 1 and Experimental procedures). Here we analysed both, ERPs, providing information about temporal aspects of WM processing, and global measures of the electric field, informing about temporal aspects of stable electric field configurations (Michel et al., 2001; Michel et al., 2004a; Murray et al., 2006).

We expected larger mean N400 amplitudes for negative compared to positive probes (Friedman and Johnson, 2000; Kutas and Federmeier, 2011). It was further hypothesised that positive probes elicit larger P300 amplitudes than negative probes. However, due to task difficulties the amplitude difference might be smaller for tonal compared to verbal stimuli. Finally, it was also hypothesised that the global field power and global dissimilarity elicited by the perception of the probe stimulus would differ between the tonal and the verbal condition (Michel et al., 2004a, also see Experimental procedures).

In contrast to the studies by Koelsch et al. (2009) and Schulze et al. (2011a, 2011b) which investigated neural correlates of rehearsal processes, the present study focused on temporal correlates of processes involving memory scanning and decision making. Therefore, the studies by Koelsch et al. and Schulze et al. on the one hand and the present study on the other, complement one another.

2.   Results

2.1.   Behavioural data

Fig. 1 displays the hit rates plotted against the false alarm rates for the tonal and verbal WM condition ($d'_{tonal} = .65$,
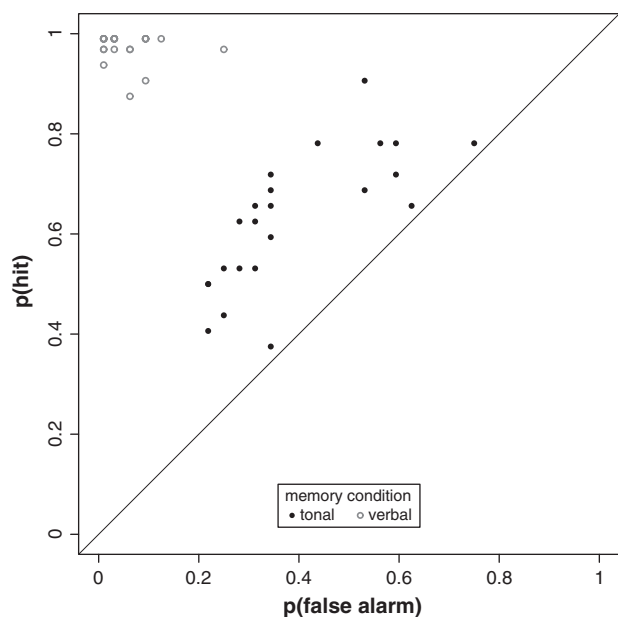
Fig. 1 – Plot of the hit-rate against the false-alarm-rate for the tonal (black filled circles) and verbal (grey open circles) memory conditions. Within conditions, each data point represents one participant.

SD=.29; d'_{verbal}=3.88, SD=.72). To assess differences between the tonal and the verbal condition a paired t-test was performed on the area A_z under the ROC-curve. In order not to violate the normal distribution assumptions (because $0<A_z<1$) scores were transformed with 2*arcsin(sqrt(A_z)) (Kirk, 1995) and the t-test was performed with the transformed scores. The analysis of the discrimination performance revealed better performance for the verbal (A_z=.99, SD=.02) compared to the tonal condition (A_z=.67, SD=.08, t_{22}=34.02, $p<.001$). In order to enable comparison with other studies, we also report the percentage of correct responses: 96% (SD=3.8%) for the verbal and 61% (SD=5.4%) for the tonal WM condition.

2.2.   ERP analysis

Because of the high performance in the verbal condition, ERPs for the tonal and the verbal condition were analysed for hits and correct rejections only. Figs. 2a and b display the ERPs for one selected electrode in each ROI. In comparison to hits, correct rejections elicited an increased N400 in the verbal memory conditions, the N400 being maximal over anterior electrodes. In the tonal condition a slightly larger (but not significant) N400 amplitude for correct rejections compared to hits can be seen only over parietal areas.

However, no dominant P300 component was identifiable in the data. Inferential statistic is therefore only reported for the N400.

To test for amplitude differences in the N400 time window (300–500 ms), a repeated-measures ANOVA with the factors memory condition (verbal, tonal), response (hit, correct rejection), hemisphere (left, right electrodes) and antpost (anterior, posterior electrodes) was conducted. The ANOVA indicated main effects for memory condition (larger mean N400 amplitude in the verbal compared to the tonal condition), response (reflecting larger N400 amplitudes for correct rejections compared to hits) and antpost (reflecting stronger negativity over anterior compared to posterior electrodes). Furthermore, the ANOVA revealed a two-way-interaction between the factors memory condition and response. Additional pairwise comparisons revealed that the main effect of response was only present in the verbal (t_{22}=4.40, $p<.001$), but not the tonal (t_{22}=0.94, $p=.36$) condition. All other effects were non-significant (see Table 1 for complete results).

Since the performance in the two memory conditions was different and thereby the number of analysed trials differed, we conducted a post-hoc power analysis to further evaluate the non-significant effect response in the tonal condition. We therefore assumed the true mean difference being of the same size in the verbal and the tonal condition but more variable in the tonal condition due to fewer correct trials. This way we estimated the effect size for the post-hoc power analysis based on the mean amplitude difference between hits and correct rejections in the verbal condition (M=1.82) and the standard deviation of the difference in the tonal condition (SD=2.41). With an effect size of d=.76, a Type I error probability of α=.05, and the given number of subjects (n=23), we computed a power estimation of (1-β)=.93. For reasons of comparison, the post-hoc power analysis for the verbal condition (post-hoc effect size d=.90, α=.05, n=23) yielded an estimation of (1-β)=.98. Thus, it is unlikely that the reduced number of valid observations in the tonal condition is responsible for the non-significant difference between hits and correct rejections.

2.3.   EEG global analysis

Fig. 3 displays the global field power (GFP) for hits and correct rejections in the tonal and verbal task. With exception of the first 150 ms, GFP was lower in the verbal compared to the tonal memory task. GFP differences between hits and correct rejections were more pronounced in the tonal compared to the verbal memory task.

The point-wise 2 (verbal vs. tonal condition)×2 (hit vs. correct rejection) repeated-measures ANOVA of the strength of the electric field revealed a stronger GFP in the tonal compared to the verbal condition from 140 to 280 ms after probe onset, and intermittent from approximately 450 ms onwards. A main effect of response lasting longer than 20 ms occurred only between 68 and 102 ms with a stronger GFP for correct rejections compared to hits. No interaction was observed.

Differences in the topography of the electric field for the tonal and verbal task were assessed separately by the global

Fig. 2 – ERPs for hits (black) and correct rejections (grey) in the (a) verbal and (b) tonal memory condition displayed for one exclusive electrode (F3, F4, P3, P4) per ROI. Probe onset is at time=0. The rectangles in the verbal condition indicate the significant ERP old/new effect of response observed in the verbal, but not the tonal condition.
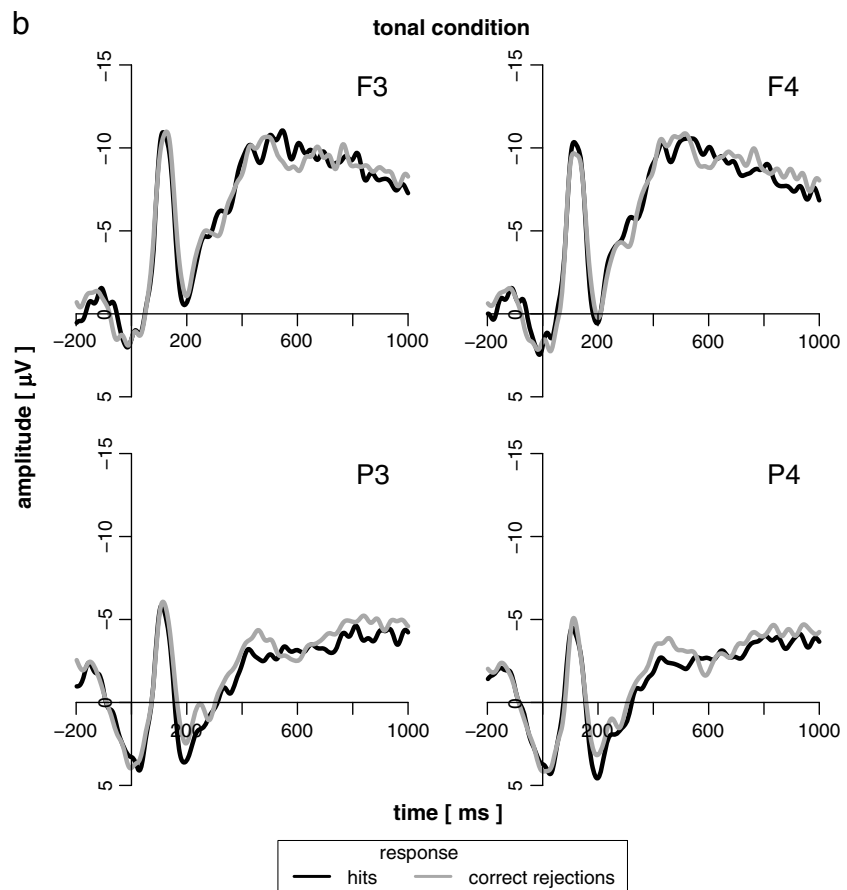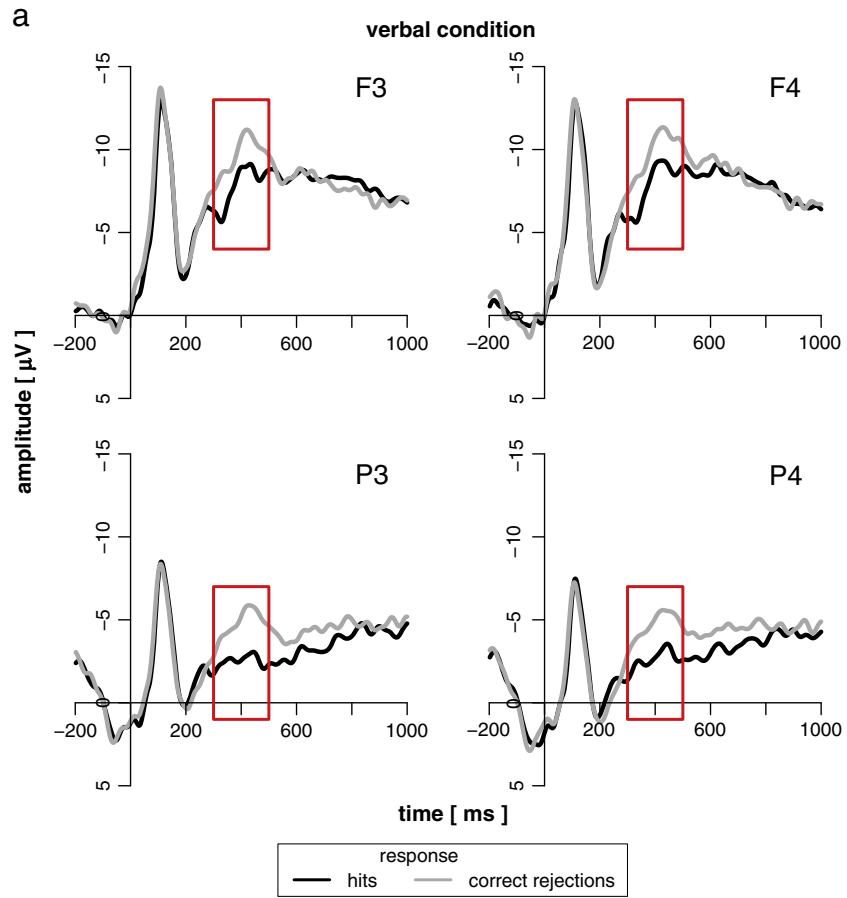
a **verbal condition**

b **tonal condition**

| Table 1 – Results of the 2 (memory condition)×2 (response)×2 (hemisphere)×2 (antpost) repeated-measures ANOVA concerning the amplitude of the N400.			
Effect	df	F	p-value
Memory condition (MC)	1,22	48.90	.002
Response (R)	1,22	12.13	.002
Hemisphere (H)	1,22	0.124	.728
Antpost (AP)	1,22	70.74	<.001
MC×R	1,22	4.96	.037
MC×H	1,22	1.25	.276
MC×AP	1,22	3.44	.077
R×H	1,22	0.02	.902
R×AP	1,22	2.79	.109
H×AP	1,22	0.08	.371
MC×R×H	1,22	1.25	.276
MC×R×AP	1,22	0.32	.576
MC×H×AP	1,22	1.42	.246
R×H×AP	1,22	3.686	.068
MC×R×H×AP	1,22	0.52	.477

Fig. 3 – Global field power for hits (black) and correct rejections (grey) in the verbal (dotted line) and tonal (normal line) memory condition as a function of time. Significant effects of memory condition and response are indicated by black horizontal lines.

dissimilarity measure (GD) for hits and correct rejections. For hits, significant differences between the tonal and the verbal condition were found for the first 60 ms, from 200 to 330 ms, 520 to 550 ms, and 620 to 640 ms after probe onset. For correct rejections differences between the tonal and the verbal condition occurred for the first 60 ms as well as between 200 and 330 ms, 468 and 494 ms, 560 and 606 ms, and 636 and 688 ms after probe onset.

The subsequent topographical analysis of the evoked potentials revealed twelve different microstates (Fig. 4) dividing the analysed time window in six segments. To identify microstates being unique to either memory condition (tonal, verbal) and/or response (hit, correct rejection), three-way repeated-measures ANOVAs with the factors memory condition, response, and microstate index were calculated for each segment (Michel et al., 2004a; Pascual-Marqui et al., 1995). Only significant effects will be reported.

During the initial phase of the probe processing (0–62 ms) different microstates occurred for the tonal and verbal condition (microstate 1 vs. microstate 2, memory condition×microstate index $F_{1,22}=21.69$, $p<.0005$). The maximum of the evoked potentials was more right lateralised and more parietal in the verbal compared to the tonal memory condition. In the next segment (63–203 ms) microstate 3 (fronto-central minimum and right lateralised parietal maximum) was dominant for all conditions, as indicated by the main effect ($F_{2,44}=7.74$, $p=.001$).

Significant two-way interactions memory condition×microstate index ($F_{2,44}=8.49$, $p=.001$) and response×microstate index ($F_{2,44}=6.49$, $p=.003$) occurred between 204 and 330 ms: microstate 4 (fronto-central minimum and right-lateralised parietal maximum) was present for hits and correct rejections in the verbal as well as for correct rejections in the tonal condition. In contrast, microstate 7 (left-lateralised frontal minimum and occipital maximum) and 8 (slightly left-lateralised frontal minimum and occipital maximum) occurred in all conditions. However, in the verbal condition both microstates differentiated hits (microstate 7) from correct rejections (microstate 8).

In the subsequent segment 331–550 ms, a main effect of microstate index occurred ($F_{2,44}=4.92$, $p=.012$; microstate 9 being present for a longer time). Additionally, a significant
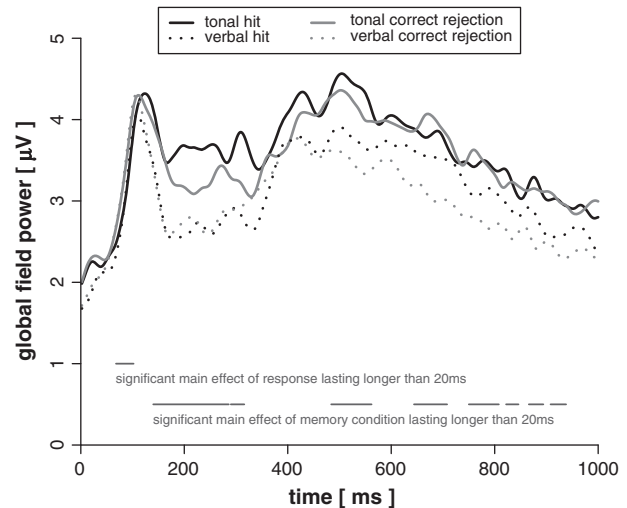
response×microstate index interaction was observed ($F_{2,44}=5.17$, $p=.009$): whereas microstate 9 (left lateralised parieto-occipital maximum) occurred for hits, microstate 10 (left lateralised parietal maximum) was present for correct rejections. To further investigate this interaction, post-hoc tests were performed. For microstates 9 and 10 in the tonal condition there was no difference between hits and correct rejections (microstate 9: $t_{22}=1.166$, $p=.256$; microstate 10: $t_{22}=-0.64$, $p=.527$). By contrast, microstate 9 and 10 clearly differentiated between hits (microstate 9) and correct rejections (microstate 10) in the verbal condition (microstate 9: $t_{22}=2.83$, $p=.009$; microstate 10: $t_{22}=-3.68$, $p=.001$).

For the following segment (551–688 ms) a significant three-way interaction memory condition×response×microstate index ($F_{2,44}=3.95$, $p=.026$) revealed the occurrence of microstate 4 in both conditions, but for microstate 9 and 11 (differing in the location of the maximum) there was a cross-over interaction for the verbal, but not the tonal condition. In the verbal condition microstate 9 was dominant for hits, whereas microstate 11 occurred for correct rejections.

In the final segment (689–1000 ms) the main effect of microstate index ($F_{3,66}=7.31$, $p<.0005$) and both two-way interactions memory condition×microstate index ($F_{3,66}=3.26$, $p=.027$) and response×microstate index ($F_{3,66}=3.32$, $p=.025$) were significant. Microstate 4 (right-lateralised maximum) was more present in the tonal, whereas microstate 12 (left-lateralised maximum) was more present in the verbal condition. Microstate 11 (fronto-central minimum and left-lateralised parietal maximum) occurred in all four conditions.

3.　Discussion

The present EEG-study aimed to investigate the temporal dynamics during a tonal and verbal working memory task. We briefly discuss the behavioural data followed by the analysis
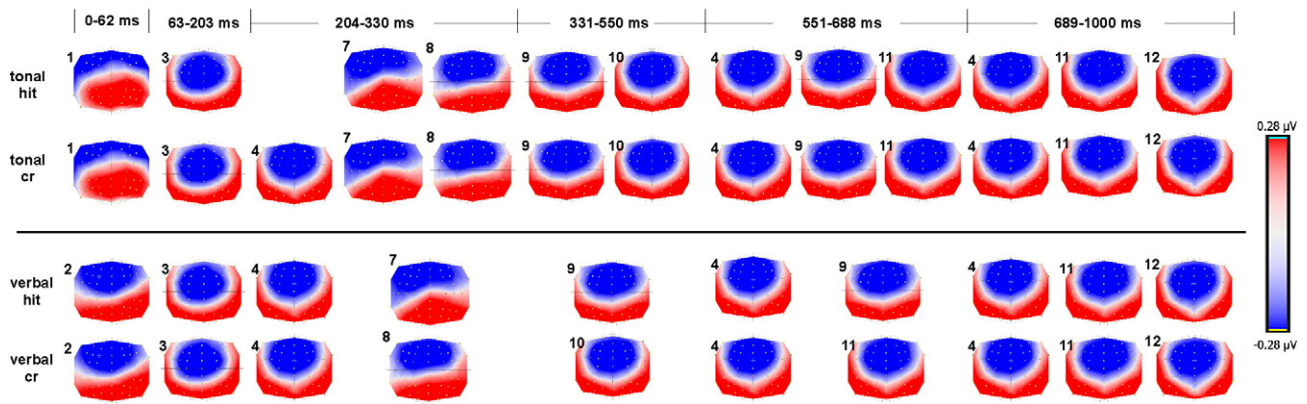
Fig. 4 – The cluster-analysis resulted in twelve different potential maps (microstates; numbered from one to twelve) shown as 2D-flat view of the scalp. Potentials range from –0.28 μV to 0.28 μV, blue indicates negative and red indicates positive potentials. Microstates are ordered in their appearance in the relevant condition. Note that even if microstates appear in both conditions they might vary in the frequency of their appearance.

of the EEG data according to local (ERP) and global (GFP, microstates) measures.

3.1. Behavioural results

Participants showed a superior performance in the verbal compared to the tonal task. The obtained percentages of correct responses (verbal: 94%, tonal: 61%) are comparable to those found by Koelsch et al. (2009, verbal: 97%, tonal: 64%) and slightly better than the nonmusicians in the study by Schulze et al. (2011a, 2011b, verbal: 85%, tonal: 56%). Although it was aimed to minimise differences in task difficulty, the tonal task was still more difficult compared to the verbal task.

3.2. ERP results

As hypothesised, we observed a stronger N400 for correctly rejected negative probes compared to hits. However, this holds only true for the verbal, but not the tonal memory condition.

For the verbal condition this is in line with the recognition literature (Friedman and Johnson, 2000; Kutas and Federmeier, 2011). The probe is compared to the actively maintained items of the memory set and negative probes (i.e. not presented in the memory set) compared to positive probes (i.e. presented in the memory set) elicited a larger N400 amplitude.

For the tonal condition no reliable N400 amplitude differences between hits and correct rejections were observable. Since we presented the same stimuli (simultaneous sine wave tone and phoneme) in the verbal and tonal task, differences due to perceptual processes can only be minimal.

Given that the amplitude difference in the verbal condition is caused by a strong signal to differentiate positive from negative probes, the lacking amplitude difference in the tonal condition suggests that no such strong signal is present. However, as indicated by the behavioural data, participants performed above chance in the tonal condition. This implies that there is some, but weak, remaining information of the memory set on which the decision is based. The microstate analysis (c.f. Section 3.4) supports this explanation, since we

observed the same two microstates in the verbal and tonal condition. By contrast, only for the verbal condition the microstates differentiated hits and correct rejections. Therefore, we suggest that participants applied a more efficient rehearsal in the verbal, but not in the tonal condition (c. f. Koelsch et al., 2009; Schulze et al., 2011a, 2011b). Hence, the information to differentiate positive from negative probes is much weaker in the tonal condition, which is reflected in the non-significant amplitude difference between hits and correct rejections.

It is necessary to remark that the number of analysed trials was smaller in the tonal compared to the verbal condition. However, the converging evidence of the behavioural data, the ERP data in conjunction with the power analysis, and the microstate analysis, corroborates the assumption that the observed differences are due to memory processes rather than statistical artefacts.

Against our expectation, no clear P300 component was observed for positive probes. Recent studies revealed a serial position effect reflected in the amplitude of the P300 which increases with recency (Chao and Knight, 1996; McEvoy et al., 1998; Patterson et al., 1991). Furthermore, long retention intervals are reflected in decreased P300 amplitudes (Chao and Knight, 1996; Kim et al., 2001; McEvoy et al., 1998; Nielsen-Bohlman and Knight, 1994). In the present study we administered a long retention interval and aggregated ERPs for positive probes over serial position. Both aspects might have led to the lack of a P300 component. However, with a modified procedure the investigation of P300 differences related to memory conditions will be fruitful for future studies.

3.3. Global field power

In line with our hypothesis, the global field power was stronger in the tonal compared to the verbal memory condition. This holds true for most parts of the analysed time window. Because differences in GFP are attributable to differences in the amount of synchronous neuronal activation and neuronal population size (Michel et al., 2001; Michel et al., 2004a) the results can be interpreted in terms of stronger engagement of

neuronal resources in the tonal compared to the verbal condition. This, in turn, reflects differences in task difficulty as supported by the behavioural data, i.e. the tonal task was more difficult and therefore resulted in a stronger activation of neural resources.

A main effect of response occurred only in an early stage of stimulus processing (68–102 ms) with a stronger GFP for correct rejections compared to hits. Positive probes (i.e. an old item) apparently lead to less neuronal activation than negative probes (i.e. new item). This indicates encoding facilitation for repeated items.

3.4.   Microstates of verbal and tonal WM

The microstate analyses revealed differences between the verbal and the tonal condition in different time segments. For the initial processing of the probe different microstates emerged for the tonal and verbal condition. In both memory conditions the physical characteristics of the probe stimuli were identical, because tones and syllables were presented simultaneously. However, the conditions were administered block-wise. Participants might have selectively attended to the task-relevant aspect of the probe. The different microstates in the tonal and verbal condition thereby reflect processing differences modulated by attention-dependent early selection (Broadbent, 1958).

In the three time windows 204–330 ms, 331–550 ms, and 551–688 ms a common pattern of microstate prevalences occurred: in the verbal condition two microstates were unique to either hits or correct rejections (7 vs. 8; 9 vs. 10; 9 vs. 11). In the tonal condition, the same microstates occurred, but did not differentiate hits and correct rejections. This complements the ERP data suggesting that the signal to differentiate hits from correct rejections is more clearly in the verbal compared to the tonal condition. The additionally occurring microstate number 4 never clearly differentiated hits from correct rejections nor tonal from verbal processing.

During the final processing window (689–1000 ms) the microstate analysis revealed differences in the predominance of the occurring states 4 and 12: microstate 4 was predominant in the tonal condition, whereas microstate 12 dominates the verbal condition. Both microstates are characterised by a fronto-central minimum amplitude, but differ in the lateralisation of the maximum amplitude: a clear right-lateralisation was observed in the tonal condition compared to a clear left-lateralisation in the verbal condition. Therefore, the right hemisphere seems to be involved more strongly during tonal a WM task, and the left hemisphere during a verbal WM task. This is in line with results by Zatorre et al. (1992) who showed in a PET study an increased activation in the left hemisphere during phonetic, but not pitch, discrimination. Other studies investigating pitch-related judgements (e.g. Gaab et al., 2003; Zatorre et al., 1994) or the detection of irregularities of musical chords (e.g. Koelsch, 2005; Koelsch et al., 2005) likewise reported a right lateralisation in the tonal conditions.

The microstate analysis is in line with the ERP and the behavioural results. Again the missing differentiation between hits and correct rejections in the tonal task is attributable to a weaker signal differentiating hits from correct rejections. In terms of evidence accumulation (Ratcliff, 1978) this suggests that in the verbal condition accumulated evidence drifts continuously to the correct barrier whereas in the tonal condition it somewhat oscillates between both boundaries. This explains the prevalence of both microstates in the tonal hit and correct rejection conditions whereas the observed cross-over interactions occur for the verbal condition.

The microstate analysis suggests that the memory scanning processes are based on a somewhat weaker representation of the memory set stimuli which results in a weaker signal differentiating positive from negative probes.

The results of the present study complement the fMRI data for nonmusicians by Schulze et al. (2011a, 2011b) who observed a larger network being activated for the verbal compared to the tonal condition, but all activated components during the tonal task were also activated during the verbal task. Different areas for tonal compared to verbal WM were observed only for musicians. If the processing in these areas corresponds to a better representation of the tonal memory set, the microstate differences between tonal and verbal stimuli (as found in the present study) should not occur. Furthermore the N400 old/new effect should emerge if the effect is independent of verbalisation (Danker et al., 2008). Future studies should investigate this by comparing musicians and nonmusicians.

3.5.   Conclusion

The present study reveals a strong difference in the N400 amplitudes for correct rejections compared to hits in the verbal, but not the tonal WM condition. This is interpreted in terms of a less efficient memory scanning process due to weak WM representation of the memory set stimuli.

The microstate analysis supports the interpretation of a weak decision signal, since different microstates emerge for hits and correct rejections in the verbal condition, but the same microstates do not differentiate between hits and correct rejections in the tonal condition.

Thus similar memory scanning processes operate in the tonal and verbal WM, but they might rely on differently pronounced representations stemming from different underlying WM networks, as shown by Schulze et al. (2011a, 2011b).

4.   Experimental procedures

4.1.   Participants

Data were obtained from 28 participants. Five participants were excluded from the analyses (four subjects responded at chance level in the tonal memory condition and one displayed extensive muscle artefacts in the EEG data). The remaining 23 participants (aged 20 to 32 years, M=24;7 years; 12 females) reported to have normal hearing, normal or corrected-to-normal vision, and no former neurological diseases. All participants were right-handed (mean laterality quotient=96.1%, range: 80–100%) according to the Edinburgh Handedness Inventory (Oldfield, 1971). Participants had normal WM capacities as indicated by normal reading span scores ranging from 2 to 6 (M=4, SD=1.26; Ems et al., 1991, German version of the reading span test by Daneman and Carpenter, 1980). None of the participants had

any special musical training in addition to the general music education in school.

4.2. Stimuli and apparatus

The stimuli were a selection of those used by Schulze et al. (2011a, 2011b), and consisted of a spoken syllable and a simultaneously presented sine wave tone (Fig. 1). A detailed description of the stimuli is provided in Schulze et al. (2011a, 2011b); in brief, the pitches of the sine wave tones corresponded to the pitches of the tones of the Western musical scale (based on A=440 Hz), ranging from 261 Hz to 523 Hz (one octave). The syllables were the names of the tones of the Western musical scale (in German, e.g. 'GIS' [gIs] (G sharp), 'C' [tse:], spoken by a professional male speaker). Each stimulus had a duration of 400 ms followed by 400 ms silence. Five stimuli were assembled as a sequence (memory set sequence) with a total duration of 3600 ms.

Recent studies comparing tonal and verbal WM (Koelsch et al., 2009; Schulze et al., 2011a, 2011b) showed that nonmusicians have greater difficulties in the tonal task (Koelsch et al. (2009) M=64% correct responses; Schulze et al. (2011a) M=56% correct responses) compared to the verbal task (Koelsch et al. (2009) M=97% correct responses; Schulze et al. (2011a) M=85% correct responses). To decrease this difference in task difficulty, only memory set sequences were presented in which all five tones of the sequence belonged to the same key, and in which three of the five tones belonged to one triad (both of these features facilitate memory performance, e.g. Koelsch and Siebel, 2005; Schmuckler, 1997). In addition, the syllables of sequences featured a high phonological similarity (e. g. 'FIS' [fIs], 'A' [a], 'H' [ha], 'CIS' [tsIs], 'DIS' [dIs]), which should exacerbate the performance for the verbal WM task (Baddeley, 2003; Conrad and Hull, 1964).

The probe consisted of a spoken syllable and a simultaneously presented sine wave tone. Within the verbal and the tonal condition, half of the probes were positive (i.e. included in the memory set sequence), counterbalanced across the five positions within the memory set sequence. A total of 16 sequences per condition with four replications per sequence were presented. The auditory stimuli were presented via loudspeakers.

Visual cues were displayed on a 15" monitor coloured on black background. The experiment was run using PRESENTATION 10.2.

4.3. Procedure

Participants were tested individually in an acoustically and electrically shielded EEG cabin. They performed four verbal and six tonal practise trials before the start of the experiment. Feedback was given only for the practise trials.

Sequences were presented pseudo randomly in a blocked design (participants started either with a verbal or a tonal block). Each experimental block started with a coloured capital letter indicating whether participants should rehearse the tones (green 'T', initial of the German word for tone) or the syllables (red 'B', initial of the German word for letter). Each trial began with the presentation of a fixation cross followed by the presentation of the memory set sequence (Fig. 5). Subsequently, participants subvocally rehearsed either the tones (tonal condition) or the syllables (verbal condition) for a duration of 16,000 ms. The pace of the rehearsal was set by a flashing cross. At the end of the rehearsal the colour of the cross changed to red or green (depending on the condition), serving as a cue for the subsequently presented probe stimulus, which was presented after 1000 ms ISI. After another delay of 2100 ms, participants were asked to indicate via a button press whether the probe stimulus was presented in the preceding memory set sequence ("old item") or not ("new item"). During the experiment, a total of ten blocks (five blocks with verbal rehearsal, and five with tonal rehearsal), each containing 13 trials were administered in a single session lasting approximately 1 h. Start condition and button allocation were counterbalanced across participants. Participants were asked to avoid eye blinking during the trials and were able to start each trial self-paced by pressing a button.

4.4. EEG data acquisition

The electroencephalogram (EEG) was recorded with Ag/AgCl electrodes from 64 electrodes placed according to the extended international 10–20 system (Jasper, 1958), referenced to the left mastoid (M1). The ground electrode was located on the sternum.
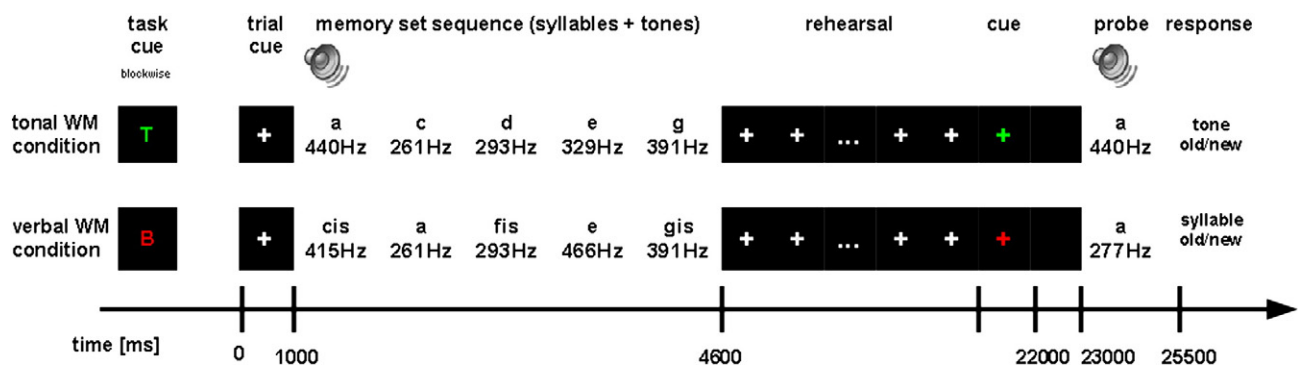

Fig. 5 – The two samples illustrate the two memory conditions. Participants listened to a memory set sequence of five stimuli (syllable+tone) and subvocally rehearsed either the syllables (verbal condition) or tones (tonal condition). When presented with a probe stimulus (syllable+tone) they had to indicate whether the tone (tonal condition) or the syllable (verbal condition) was presented in the memory set sequence.

To control for eye movement artefacts, bipolar horizontal and vertical electrooculograms (EOGs) were recorded. Signals were amplified with two synchronised PORT-32/MREFA amplifiers (Twente Medical System BV, Enschede, Netherlands) and digitised with a sampling rate of 500 Hz.

After data acquisition, EEG-data were re-referenced offline to the arithmetic mean of the mastoid electrodes (M1 and M2) to obtain a symmetric reference, and bandpass-filtered between 0.4 Hz and 25 Hz. The extracted probe stimulus epochs ranged from 200 ms before to 1000 ms after probe onset. Oculomotor and other artefacts were eliminated by rejecting sampling points exceeding 50 μV at any EEG electrode. Within participants epochs were averaged for each memory condition (verbal, tonal) and for each classified response condition (hit, correct rejection). Within each condition on average four trials were rejected due to artefacts. Out of 32 administered trials the average number of clean trials per memory and response condition was 17.5 (SD=5.3) for tonal hits, 16.6 (SD=5.1) for tonal correct rejections, 26.6 (SD=6.4) for verbal hits, and 26.4 (SD=4.8) for verbal correct rejections.

4.5.   Data analysis

4.5.1.   Behavioural data analysis
According to the signal detection theory, responses were classified into hits, misses, false alarms and correct rejections. The area under the ROC-curve A_z was taken as measure of discrimination performance within each condition.

4.5.2.   EEG data analysis
EEG data analysis was conducted with a multi-step procedure referred to as electrical neuroimaging (Michel et al., 2001; Michel et al., 2004a). It comprises local as well as global measures of the electric field at the scalp. The standard procedure and its underlying rationale are described in detail elsewhere (e.g. Michel et al., 2001; Michel et al., 2004a; Murray et al., 2006). The analysis was performed using the Cartool software by Denis Brunet (see Acknowledgements).

ERPs were computed applying a baseline correction for the pre-stimulus window of 200 ms. To test lateralisation and topography effects of the ERPs between conditions, four regions of interest (ROI) were defined: left-anterior (AF7, AF3, F5, F3, FC5, FC3), right-anterior (AF4, AF8, F4, F6, FC4, FC6), left-posterior (CP5, CP3, P5, P3, PO7, PO3) and right-posterior (CP4, CP6, P4, P6, PO4, PO8). For the N400 analysis a time window from 300 to 500 ms was selected. P300 effects were inspected in a time window from 500 to 750 ms.

A data driven spatio-temporal analysis was performed to assess changes in the response strength and the response topography of the electric field of the scalp using the global field power (GFP) and the global dissimilarity (GD) respectively (Lehmann and Skrandies, 1980; Michel et al., 2004b; Skrandies, 1990). The GFP can be understood as the spatial standard deviation of the electric field at the scalp, whereas the GD serves as a measure for the difference between the configurations of two subsequent electric fields. Different configurations are indicative of different active sources in the brain (Michel et al., 2004a).

For this analysis the EEG data was re-calculated to the common average reference (Michel et al., 2004a). Differences in the GFP between memory conditions and/or response classifications were assessed for each electrode and time point. A significance level of $p=.01$ (without further correction for multiple comparisons, which is in accordance with the literature, c.f. Lehmann and Skrandies, 1980; Michel et al., 2004b; Skrandies, 1990) and a minimum duration of 20 ms (corresponding to 10 data points) was applied. Timepoint-wise 2 (verbal vs. tonal condition)×2 (hit vs. correct rejection) repeated measurement ANOVAs were conducted concerning differences in the GFP. A non-parametric bootstrapping method (Murray et al., 2006; Srebro, 1996) was applied to assess differences in the GD between hits and correct rejections in the verbal and tonal memory condition. In addition, a topographical pattern analysis was performed to identify successive functional microstates. This comprises a K-means cluster analysis (Pascual-Marqui et al., 1995) to identify most dominant potential maps within each condition and a spatial correlation procedure to determine microstates which are specific for a given condition. It results in (a) data driven segmentation of the analysed time window in sub-windows and (b) a measure of relative microstate presence which was in turn submitted to a 2 (verbal vs. tonal condition)×2 (hit vs. correct rejection)×i (microstate index) repeated measure ANOVA. Significant interactions served as indicator for maps being unique to a certain condition in a specified time window (Michel et al., 2004a; Pascual-Marqui et al., 1995; Santis et al., 2007).

REFERENCES

Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: a proposed system and its control processes. In: Spence, K.W., Spence, J.T. (Eds.), The Psychology of Learning and Motivation: Advances in Research and Theory. Academic Press, New York, pp. 89–195.

Atkinson, R.C., Shiffrin, R.M., 1971. The control of short-term memory. Sci. Am. 225, 82–90.

Baddeley, A.D., 1992. Working memory. Science 255, 556–559.

Baddeley, A.D., 2003. Working memory: looking back and looking forward. Nat. Rev. Neurosci. 4, 829–839.

Baddeley, A.D., Hitch, G.J., 1974. Working memory. In: Bower, G. (Ed.), The Psychology of Learning and Motivation, vol. 8. New York, Academic Press, pp. 47–89.

Broadbent, D.E., 1958. Perception and Communication. Pergamon Press, London.

Chan, A.S., Ho, Y.C., Cheung, M.C., 1998. Music training improves verbal memory. Nature 396, 128.

Chao, L.L., Knight, R.T., 1996. Prefrontal and posterior cortical activation during auditory working memory. Cogn. Brain Res. 4, 27–37.

Chao, L.L., Nielsen-Bohlman, L., Knight, R.T., 1995. Auditory event-related potentials dissociate early and late memory processes. Electroencephalogr. Clin. Neurophysiol. 96, 157–168.

Conrad, R., Hull, A.J., 1964. Information, acoustic confusion and memory span. Br. J. Psychol. 55, 429–432.

Crites, S.L., Devine, J.V., Lozano, D.I., Moreno, S., 1998. Event-related potentials and serial position effects in a visual probe recognition task. Psychophysiology 35, 293–304.

Crites, S.L., Delgado, P., Devine, J.V., Lozano, D.I., 2000. Immediate and delayed stimulus repetitions evoke different ERPs in a serial-probe recognition task. Psychophysiology 37, 850–858.

Curran, T., 2004. Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. Neuropsychologia 42, 1088–1106.

Daneman, M., Carpenter, P.A., 1980. Individual differences in working memory and reading. J. Verb Learn Verb Be 450–466.

Danker, J.F., Hwang, G.M., Gauthier, L., Geller, A., Kahana, M.J., Sekuler, R., 2008. Characterizing the ERP old–new effect in a short-term memory task. Psychophysiology 45, 784–793.

Deutsch, D., 1970. Tones and numbers: specificity of interference in immediate memory. Science 168, 1604–1605.

Ems, M., Troppmann, N., Schuri, U., 1991. Reading Span — Deutsche Version. EKN- Materialien für die Rehabilitation, München.

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

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., Schlaug, G., 2003. Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. Neuroimage 19, 1417–1426.

Golob, E.J., Starr, A., 2004. Serial position effects in auditory event-related potentials during working memory retrieval. J. Cogn. Neurosci. 16, 40–52.

Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area spt. J. Cogn. Neurosci. 15, 673–682.

Jasper, H., 1958. Report of the committee on methods of clinical examination in electroencephalography. Electroen. Clin. Neuro. 10, 370–371.

Kim, M., Kim, J., Kwon, J.S., 2001. The effect of immediate and delayed word repetition on event-related potential in a continuous recognition task. Cogn. Brain Res. 11, 387–396.

Kirk, R.E., 1995. Experimental Design: Procedures for the Behavioral Sciences, 3 ed. Brooks/ Cole, Pacific Groove.

Koelsch, S., 2005. Neural substrates of processing syntax and semantics in music. Curr. Opin. Neurobiol. 15, 207–212.

Koelsch, S., Siebel, W.A., 2005. Towards a neural basis of music perception. Trends Cogn. Sci. 9, 578–584.

Koelsch, S., Gunter, T.C., Wittfoth, M., Sammler, D., 2005. Interaction between syntax processing in language and in music: an ERP study. J. Cogn. Neurosci. 17, 1565–1577.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., Gruber, O., 2009. Functional architecture of verbal and tonal working memory: an FMRI study. Hum. Brain Mapp. 30, 859–873.

Kok, A., 2001. On the utility of P3 amplitude as a measure of processing capacity. Psychophysiology 38, 557–577.

Kutas, M., Federmeier, D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Lee, K.M., 2002. Does a separate pitch memory system exist? Evidence from interference in memory for tonal pitch. In: Stevens, C., Burnham, D., McPherson, G., Schubert, E., Renwick, J. (Eds.), Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney 2002. Causal Productions, Adelaide, pp. 687–690.

Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48, 609–621.

McEvoy, L.K., Smith, M.E., Gevins, A., 1998. Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. Cereb. Cortex 8, 563–574.

Michel, C.M., Thut, G., Morand, S., Khateb, A., Pegna, A.J., Peralta, R.G., 2001. Electric source imaging of human brain functions. Brain Res. Brain Res. Rev. 36, 108–118.

Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Peralta, R.G., 2004a. EEG source imaging. Clin. Neurophysiol. 115, 2195–2222.

Michel, C.M., Seeck, M., Murray, M.M., 2004b. The speed of visual cognition. Clin. Neurophysiol. Suppl. 57, 617–627.

Mondor, T.A., Morin, S.R., 2004. Primacy, recency, and suffix effects in auditory short-term memory for pure tones: evidence from a probe recognition paradigm. Can. J. Exp. Psychol. 58, 206–219.

Murray, M.M., Imber, M., Javitt, D., Foxe, J., 2006. Boundary completion is automatic and dissociable from shape discrimination. J. Neurosci. 15, 12043–12054.

Nielsen-Bohlman, L., Knight, R., 1994. Event-related potentials dissociate immediate and delayed memory. In: Heinze, T.M.H.J., Mangun, G. (Eds.), Cognitive Electrophysiology: Basic and Clinical Research. Birkhauser, Boston, pp. 169–183.

Oldfield, R., 1971. The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia 9, 97–113.

Pascual-Marqui, R., Michel, C., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42, 658–665.

Patterson, J.V., Pratt, H., Starr, A., 1991. Event-related potential correlates of the serial position effect in short-term memory. Electroencephalogr. Clin. Neurophysiol. 78, 424–437.

Pechmann, T., Mohr, G., 1992. Interference in memory for tonal pitch: implications for a working-memory model. Mem. Cognit. 20, 314–320.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118, 2128–2148.

Ratcliff, R., 1978. A theory of memory retrieval. Psych. Rev. 85, 59–108.

Rugg, M.D., Coles, M.G.H. (Eds.), 1995. Electrophysiology of Mind: Event-related Potentials and Cognition. Oxford University Press, New York.

Salamé, P., Baddeley, A.D., 1989. Effects of background music of phonological short-term memory. Q. J. Exp. Psychol. 41a, 107–122.

Santis, L.D., Clarke, S., Murray, M.M., 2007. Automatic and intrinsic auditory "what" and "where" processing in humans revealed by electrical neuroimaging. Cereb. Cortex 17, 9–17.

Schendel, Z.A., Palmer, C., 2007. Suppression effects on musical and verbal memory. Mem. Cognit. 35, 640–650.

Schmuckler, M.A., 1997. Expectancy effects in memory for melodies. Can. J. Exp. Psychol. 51, 292–306.

Schulze, K., Zysset, S., Mueller, K., Friederici, A.D., Koelsch, S., 2011a. Neuroarchitecture of verbal and tonal working memory in non-musicians and musicians. Hum. Brain Mapp. 32, 771–783.

Schulze, K., Mueller, K., Koelsch, S., 2011b. Neural correlates of strategy use during auditory working memory in musicians and non-musicians. Eur. J. Neurosci. 33, 189–196.

Semal, C., Demany, L., Ueda, K., Hallé, P.A., 1996. Speech versus nonspeech in pitch memory. J. Acoust. Soc. Am. 100, 1132–1140.

Skrandies, W., 1990. Global field power and topographic similarity. Brain Topogr. 3, 137–141.

Srebro, R., 1996. A bootstrap method to compare the shapes of two scalp fields. Electroencephalogr. Clin. Neurophysiol. 100, 25–32.

Sternberg, S., 1966. High speed scanning in human memory. Science 153, 652–654.

Williamson, V.J., Baddeley, A.D., Hitch, G.J., 2010. Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. Mem. Cognit. 38, 163–175.

Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256, 846–849.

Zatorre, R.J., Evans, A.C., Meyer, E., 1994. Neural mechanisms underlying melodic perception and memory for pitch. J. Neurosci. 14, 1908–1919.