Below is the unedited draft of the article that has been accepted for publication
(© Cognitive Processing, 2007, V. 8. No 3. P. 183-199)

Composition of Brain Oscillations and Their Functions in the Maintenance of Auditory, Visual and Audio-Visual Speech Percepts: An Exploratory Study

Alexander A. Fingelkurts,[a*] Andrew A. Fingelkurts,[a] Christina M. Krause[b]

[a] BM-SCIENCE – Brain & Mind Technologies Research Centre,
P.O. Box 77, FI-02601 Espoo, Finland.
[b] Cognitive Science Unit, Department of Psychology, University of Helsinki
PB 9, 00014 University of Helsinki, Finland.

Abstract: In the present exploratory study based on 7 subjects, we examined the composition of magnetoencephalographic (MEG) brain oscillations induced by the presentation of an auditory, visual, and audio-visual stimulus (a talking face) using an oddball paradigm. The composition of brain oscillations were assessed here by analyzing the probability-classification of short-term MEG spectral patterns. The probability index for particular brain oscillations being elicited was dependent on the type and the modality of the sensory percept. The maintenance of the integrated audio-visual percept was accompanied by the unique composition of distributed brain oscillations typical of auditory and visual modality, and the contribution of brain oscillations characteristic for visual modality was dominant. Oscillations around 20 Hz were characteristic for the maintenance of integrated audio-visual percept. Identifying the actual composition of brain oscillations allowed us (a) to distinguish two subjectively/consciously identical mental percepts, and (b) to characterize the types of brain functions involved in the maintenance of the multi-sensory percept.

Keywords: Audio-visual percept maintenance; McGurk effect; Short-term spectral patterns; Theta, alpha, beta oscillations; MEG; Adaptive classification; Brain functions.

1. Introduction

Since our perception of the world is multimodal in nature, knowledge of the underlying neurophysiology would seem very important in understanding of human brain functions (for reviews, see Calvert 2001; King and Calvert 2001; Ernst and Bulthoff 2004). Perceptual integration can be defined as a psychophysiological process that not only facilitates detection by amplification of the sensory signals but also can combine these signals to form a new, multimodal representation of the source object (O'Hare 1991). The neural representation of a multimodal sensory event represents a perceptual brain state which causes the activation of a certain category in the mind, i.e., the percept. A conscious percept (the subjective experience that one is aware of) within one modality can be dramatically affected by stimulation within another. For example, when a syllable (auditory /ba/ + visual /ga/) is presented, the subjects usually perceive the stimulus as /da/ (McGurk and MacDonald 1976). This so-called "McGurk" effect demonstrates that visual information strongly affects auditory perception.

Even though the ability of the brain to sustain neural activity in the absence of stimuli has been extensively demonstrated on both the single neuron (Fuster and Alexander 1973; Funahashi et al. 1989) and the neural population (Courtney et al. 1997; Zarahn et al. 1999) levels, the neural mechanisms underlying the maintenance of a stimulus representation remain poorly understood (Gazzaley et al. 2004). Maintenance can be defined as persistent neural activity in the absence of stimulus in the environment. To our best knowledge no study has been undertaken which focuses on the actual neural mechanisms underlying the maintenance of a multimodal representation of the sensory event.

In order to assess the maintenance of a conscious multimodal percept of a sensory event, it is important to determine the time frame for the analysis. For a percept to be sustained despite the removal of the stimulus, there must be a period of time in which the representation of the percept is dissociated from any sensory input (Large et al. 2005). For a percept to be conscious, time up to 500 msec is necessary to elicit the conscious experience of an event (Libet 1993; Taylor and McCloskey 1990). Indeed, the transition between the psychological detection of a sensory signal without awareness and the detection with awareness, has been found to be controlled simply by differences in the duration of repetitive ascending activations of the sensory cortex, with a minimum duration of up to approximately 500 msec (Libet 1993). Hence, the formation of a conscious and integrated multimodal representation of a non-brief dynamic sensory stimulus is completed after stimulus offset during the post-stimulus interval (Mesulam 1998; Näätänen and Winkler 1999). Considering (1) that it has been demonstrated behaviourally that the reproduction of visual or auditory stimuli of different duration can be done veridically and with a small variance up to approximately 2 to 3 seconds, but not beyond (Pöppel 1994, 2004; Mukamel et al.

2004), (2) that neural assembly holds its activity after activation in the range of several tens of milliseconds to several hundreds of milliseconds (Varela 1995), (3) that some researchers refer to long, slow induced effects accompanying higher cognitive processing (Kutas and Hillyard, 1980; Friederici 2002; see Lebib et al. 2004, for the "McGurk effect"), and (4) that the persistence found in brain activation correlates with the participants' phenomenological experience of perceptual persistence (Large et al., 2005), then any time interval in a 2 sec period should reflect the maintenance of stimulus representation. However, a shorter period (of up to several hundred milliseconds) should be used in order to minimize the interrelationships between the neural representation and internal brain processes (Usher and Niebur 1997). Thus, the maintenance of stimulus representation may be reflected in the induced post-stimulus brain activity. This activity, which corresponds to the information content of the stimulus perception and sensory memory, may signify a stage when the outputs of the central afferent processes are integrated and mapped on the neurophysiological substrate of sensory memory (Näätänen and Winkler 1999).

However, in the period of the maintenance of stimulus representation there are several task-unspecific and task-specific processes going on in parallel. These processes are arousal, attention, encoding, sensory short-term memory or simply the anticipation of making a response. While the contribution of the most unspecific processes (such as arousal, anticipation of making a response and so on) can be easily controlled in the experiment (see below and Section 2.6), it is difficult to disentangle maintaining process from other task-specific processes such as attention, encoding, and sensory short-term memory. At the same time, these task-specific processes appear to work in concert to subserve the maintenance of representations (Mesulam 1998; Näätänen and Winkler 1999; McCollough et al., 2007; see also Basar et al. 2004). As such, for the purpose of this paper by the 'maintenance' we will understand a mixture of task-specific processes such as attention, encoding, and sensory short-term memory in the post-stimulus interval which is not exceeded a 2 sec period.

In order to assess the maintenance of a conscious multimodal percept (a stimulus after-effect distributed in the neural basis of sensory memory), it seems reasonable to examine the post-stimulus EEG/MEG intervals (free from the direct influences of any event-related potentials/fields (ERP/ERF)). However, identifying components with very long latencies (more than 400 ms) is technically difficult because their effects become submerged in the ongoing EEG/MEG field complex. At the same time, estimations of short-term spectral patterns (SP) – that are induced by stimulus events and based on Short-Time FFT – provide an adequate measure for EEG/MEG oscillatory patterns with long latencies (Fingelkurts et al. 2002; 2003b).

Since oscillatory activity is intrinsic to the functioning of the brain, it is possible to suppose that information about a stimulus is coded on the one hand, by oscillation frequency and on the other hand, by the spatial distribution of oscillatory systems (Singer, 1994). It has been proposed that brain oscillatory systems act as possible communication networks with functional relationships to the integrative brain functions (Basar et al. 2001a). It is assumed that brain oscillations are of fundamental importance for mediating and distributing “higher-level” processes in the human brain (Klimesch 1999; Basar et al. 2001c). However, there are no studies on the representation of a multimodal percept by brain oscillations and only a few studies have investigated the role of brain oscillations during human multisensory information processing (Sakowitz et al. 2000, 2001; Krause et al. 2001). However, in all of these studies averaging procedures were used and the frequency bands were predefined and taken in isolation from each other. This prevents researchers from obtaining the actual/natural composition of brain oscillations involved in the process of multimodal integration. At the same time, it was demonstrated that multiple frequencies are generated simultaneously in the induced EEG/MEG and that their superposition has important functional implications (Basar et al. 2004).

Thus, the aim of this study is to explore in detail the role of brain oscillations within a broad frequency range (3-30 Hz) in relation to the maintenance of a conscious and integrated multisensory percept in an audio-visual oddball task with both congruent and incongruent (McGurk type) deviant audio-visual stimuli. As there is extensive data on the frequency-dependent functional significance of oscillatory brain activity (Lopes da Silva et al. 1980; Arroyo et al. 1993; Klimesch 1996, 1999; Pfurtscheller et al. 1996; Bullock 1997; Doppelmayr et al. 1998; Basar 1999; Pulvermuller et al. 1999; Basar et al. 2001a,b,c; to mention just a few), a frequency domain approach should reveal which types of brain functions are involved in the maintenance of an audio-visual integrative percept. It is of further interest whether any typical/steady composition of brain oscillatory patterns induced by an acoustic, visual and audio-visual stimulation exist and which frequencies are included in these patterns. The exact compositions of brain oscillations were assessed here by a probability-classification analysis of the short-term MEG spectral patterns (suggested by Kaplan et al. 1999; details see in Fingelkurts et al. 2003a) in the totality of post-stimulus intervals categorized with respect to the preceding stimulus. Supposing that brain oscillations induced by the same stimulus events are systematic and that unspecific/unconscious neurophysiologic processes that take place in the same time window are random during each trial and between the subjects, then we would estimate mostly the dependence of the probability of occurrence of induced spectral patterns (SPs) of a particular type, which reflect ‘internal’ events (sensory percepts) on repetitive ‘external’ events – such as the presentation of a stimulus. By not averaging the post-stimulus intervals, we were able to examine in detail the actual composition of brain oscillations in association with the maintaining of a multimodal percept (for discussion on the problems of averaging techniques see Laskaris and Ioannides 2001; Fingelkurts et al. 2002). The experiment was conducted in the audio-visual, visual and auditory stimulus modalities. The mismatch fields (Möttönen et al. 2002) and cortical operational synchrony (Fingelkurts et al. 2003c) analyzed from the same data have already been published. The present study substantially extends the previous two in that: (a) in contrast to the work of Möttönen et al. (2002), the present study examines post-stimulus MEG intervals (free from the direct influences of event-related potentials/fields (ERP/ERF)) and focuses on induced long-latency (more than 400 ms) events which most likely reflect the maintenance of uni- and multimodally determined percepts; (b) by investigating the local processes in the brain cortex (analysis here was done separately for each MEG location), the present study adds to the work of Fingelkurts et al. (2003c) where the authors examined functional connectivity between different cortex locations during multimodal speech integration.

2. Material and methods

2.1. Subjects

Ten Finnish-speaking volunteering subjects (three females, one left-handed, mean age: 28±5.1) participated in the experiment. All subjects had normal hearing and vision (self-reported). The data from one male subject was excluded from further analysis because of extensive artefacts in the recordings. Two subjects did not have the McGurk effect and these were also omitted from further analyses. Seven subjects had the McGurk illusion. These seven subjects (mean age:

29±5.3) also participated in the visual experiment and three of them (mean age: 26±3.2) participated in the auditory experiment.

2.2. Stimuli

Meaningless disyllables (vowel-consonant-vowel) uttered by a female Finnish speaker were recorded in a chamber with a professional video camera. The speaker was requested to maintain the same intonation, tempo, and vocal intensity whilst producing the sounds and was also told to avoid blinking during her articulation. The speaker was recorded with lights focused on her face such that minimal shadowing occurred. In the recorded image, the speaker's entire head was visible against a black background. Visual clips (frame rate 25 Hz) and sound files (digitized at 22 050 Hz) were extracted from the digital video for each stimulus utterance ("ipi", "iti", and "ivi"). The visual utterances were about 800 ms in length. The duration of the acoustic /ipi/ was 588 ms and that of the acoustic /iti/ was 581 ms. The action of opening the mouth in the visual /ipi/ and /iti/ stimuli began 230 ms prior to the acoustic utterance onset, when presented together. The total lengths of acoustic and visual stimulus files were 1640 ms including the periods of silence where the face had a closed mouth before and after the utterances. The periods of silence where the mouth of the face was closed after stimulus offset and before the onset of the following stimulus comprised the inter-stimulus interval (880 ms). The audio-visual experiment included four stimuli: congruent "ipi" (auditory /ipi/ + visual /ipi/), congruent "iti" (auditory /iti/ + visual /iti/), incongruent "iti" (auditory /ipi/ + visual /iti/) and congruent "ivi" (auditory /ivi/ + visual /ivi/). The visual experiment contained only the visual parts of these stimuli and the auditory experiment contained only the auditory parts.

2.3. Stimuli Presentation

The stimulus sequences were presented to the subjects using the "Presentation" software (Neurobehavioral Systems, Inc, 2001). The audio-visual stimuli consisted of frequent (p = 0.85) congruent "ipi" stimuli (*standards*) and infrequent *deviant* congruent (p = 0.05) and incongruent (p = 0.05) "iti" stimuli. Deviant congruent "ivi" stimuli were presented as targets (p = 0.05) which the subjects were instructed to count silently during the registration in order to check that the subjects were consciously paying attention to the stimuli. The auditory stimuli were delivered binaurally to the subjects through plastic tubes and earpieces. The intensity of the sound was adjusted to 55 dB above the subject’s hearing threshold (defined for the audio-visual stimulus sequence). The visual stimuli were projected into the measurement room through a data projector. The height of the face stimulus was 12 cm and projected at a distance of 105 cm from the subject.

In the unimodal-stimuli audio-only and visual-only experiments, the visual stimuli and the audio stimuli respectively were not presented. However, these experiments were in all other aspects identical to the audio-visual experiment.

2.4. Procedure

The audio-visual experiment consisted of 3-4 sessions, each lasting between 15-20 minutes. The subjects were instructed to concentrate on the stimuli and silently count the number of “ivi” utterances. After each session the subjects were asked to report their counting. This procedure ensured that the subjects followed the instructions correctly. In order to assess how the subjects perceived the incongruent audio-visual (McGurk-type) utterances, a behavioural test was carried out during one of the breaks between the experimental sessions. In this test a sequence consisting of 12 incongruent deviants, 6 congruent deviants, 12 targets and 94 standards was presented. The subjects were instructed to repeat each utterance aloud immediately after identifying what they *heard*. The experimenter wrote down the responses. All subjects identified perfectly the congruent deviants, standards and targets. Seven subjects perceived the incongruent deviants always as “iti” (demonstrating that these subjects had the McGurk effect). These subjects could not distinguish the congruent “iti” (auditory /iti/ + visual /iti/) from the incongruent “iti” (auditory /ipi/ + visual /iti/) stimuli, thus proving the presence of two subjectively identical mental percepts of two different stimuli. Two subjects reported frequently hearing “ipi”, when incongruent deviants were presented (demonstrating that these subjects did not have the McGurk effect). These two subjects were excluded from the study.

The visual and auditory experiment consisted of two sessions each lasting 15-20 minutes. The task was to count silently “ivi” utterances and to report their counting after each session. There was always an interval of at least one week between the audio-visual, visual and auditory experiments. The order of these experiments was counterbalanced across subjects.

2.5. MEG recording

The magnetoencephalogram (MEG) was recorded continuously in a magnetically shielded room with a 306-channel whole-head device (Neuromag Vectorview, Helsinki, Finland). The device has two sensor elements; two orthogonal planar gradiometers and one magnetometer.

The data was digitized at 300 Hz. The passband of the MEG recordings was 0.06-100 Hz. Approximately 100 responses from the subjects to each deviant stimulus and approximately 2000 responses to the standard stimuli were collected. Neuromagnetic activity was recorded continuously and stored on an optical disc for later analysis off-line. Epochs containing artefacts due to eye blinks, significant muscle activity, or movements were automatically rejected. The presence of an adequate signal was determined by visually checking each raw signal on the computer screen after automatic artefact rejection.

2.6. Data processing

We examined the post-stimulus MEG data (still face, no sound) extracted shortly (40 ms) after the stimulus offset. This 40 ms delay enabled us to exclude any possible evoked response, which can be triggered by the offset of the stimulus. Hence, the extracted post-stimulus intervals were not influenced by any artifact from the stimulus-events themselves. By analyzing the event-related field (ERF) (Möttönen et al. 2002) and event-related (de)synchronization (ERD/ERS) (Krause et al. 2001) using the same data, it was demonstrated that the majority of MEG changes associated with ERF and ERD/ERS were completed around 400 ms after the stimulus onset, meaning that they were completed by the time the post-stimulus intervals in the present study were extracted. A procedure was applied to “cut” 840 ms data segments in order to extract data categorized with respect to the preceding stimulus (belonging to post-standard, or post-deviant-congruent, or post-deviant-incongruent intervals). Thereafter, the data-segments for each stimulus type were “glued” together (Fig. 1).

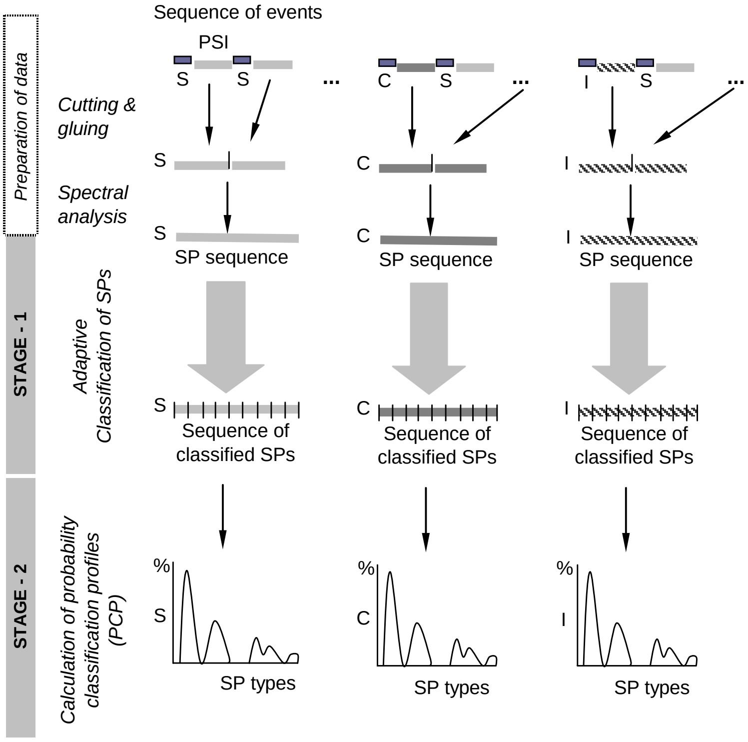


Fig. 1. The scheme of the data processing. Extractions of the corresponding post-stimulus intervals (still face, no sound) were done shortly (40 ms) after stimulus offset. SP – spectral pattern, PCP – the probability classification profile, PSI – post-stimulus interval, S – standard stimuli, C – deviant-congruent stimuli, I – deviant-incongruent stimuli.

Applying this procedure resulted in a sequence of concatenated MEG data, sorted stimulus-wise. Short-term non-stationarities in the neighbourhood of the connection points were eliminated by searching for identical groups of data-points (typically three or four) in each data set, and 'stitching' the two data sets together at these points. A forty millisecond delay for the post-stimulus MEG data extraction (see above) enabled us to search for match data-points.

As a result, the complete MEG data-streams were split into three distinct segments for audio-visual (AV) conditions: AV(S) for standard stimuli, AV(C) for deviant congruents, and AV(I) for deviant incongruents; and into two distinct segments for auditory (A) and visual (V) conditions: A(S), V(S) for standard stimuli, A(D), V(D) for deviant stimuli respectively.

Due to the technical requirements of the tools later used to process the data, with a converted sampling rate of 128 Hz, 64 MEG locations (Fig. 2, A) were analyzed which roughly correspond to the standard EEG sites (extended EEG 10-20 International system, Nuwer et al., 1998): $AF_{7/8}$, $AF_{3/4}$ (4 MEG locations), AFz, $F_{9/10}$, $F_{7/8}$, Fz, $F_{5/6}$, $F_{3/4}$, $FT_{9/10}$, $FT_{7/8}$, $FC_{5/6}$, $FC_{3/4}$, $T_{5/6}$ (4 MEG locations), $T_{3/4}$ (4 MEG locations), $CP_{5/6}$, $CP_{3/4}$, $CP_{1/2}$, CPz (2 MEG locations), $C_{5/6}$, Cz (2 MEG locations), $C_{3/4}$, $C_{1/2}$, Pz (2 MEG locations), $P_{5/6}$, $P_{3/4}$, $P_{1/2}$, Oz (2 MEG locations), $O_{1/2}$ (6 MEG locations). One may think that selection of a subset of the MEG locations, which are equivalent to the EEG 10-20 montage, would be a source of error because such an approach does not take into account variances in head sizes. This is unlikely because the inter-individual variability of the results obtained (see below) was very low.

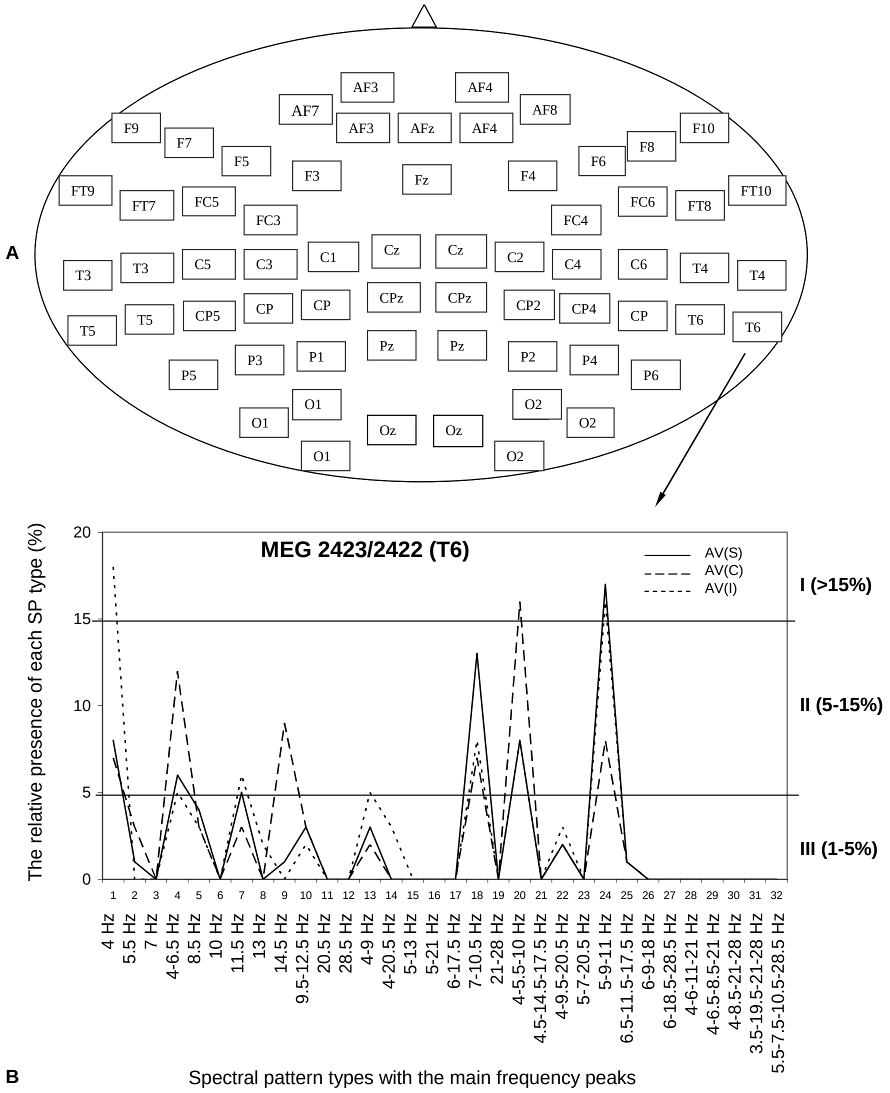


Fig. 2. (A) Positions of MEG sensors on the head surface with corresponding EEG labels (extended 10-20 International system). (B) Example of the probability classification profiles for MEG 2423/2422 location (corresponds to T6 in EEG).
AV(S) – audio-visual standard stimuli; AV(C) – audio-visual deviant-congruent stimuli; AV(I) – audio-visual deviant-incongruent stimuli. For (B): The X-axis displays the labels (numbers) of spectral pattern types from 1 to 32 and their main frequency peaks. The Y-axis displays the share of the corresponding spectral patterns in the percentage from the total number of the classified spectral patterns. A line graphic was chosen instead of a bar for the ease of comparison. (Note that X-axis consists of 32 discrete values, all the in-between values are meaningless).

Prior to the spectral analysis, each MEG data sequence (corresponding to different stimulus conditions) was bandpass filtered in the 3-30 frequency range. This frequency range was chosen because approximately 98% of spectral power lies within it (Thatcher, 2001) and it has been demonstrated that brain oscillations within this frequency range might respond to the perception of audio-visual information (Sakowitz et al. 2000; Krause et al. 2001). Since inter-stimulus intervals in the present study were constant, high-pass filtering the data with a 3 Hz cut-off frequency was chosen to reduce the amplitudes of possible anticipatory potentials (Teder-Sälejårvi et al. 2002). The gamma frequency range was not included in the analysis because the pilot study of the same data (Krause et al. 2001) in which a wide range 0.3-45 Hz frequency was analyzed, did not reveal any changes in the gamma frequency band.

Thereafter, the individual power spectra were calculated in the range of 3-30 Hz with a 0.5 Hz resolution (55 values), using FFT with a 2-sec Hanning window shifted by 50 samples (0.39 s) for each of the two gradiometers (Fig. 1). These values proved the most effective for revealing oscillatory patterns from the signal based on previous studies (Levy 1987; Kaplan 1998, for details, see Fingelkurts et al. 2006b). A sliding spectral analysis is adequate for the purpose of the present study since the focus of our interest has not been the characterization of each individual post-stimulus interval separately but rather the search for actual brain oscillations in the totality of the post-stimulus intervals. In other words, it was important to assess the general statistical regularities (in SP description) for the totality of post-stimulus intervals categorized with respect to the preceding stimulus. Additionally, sliding spectral analysis (1) reduces the random error of the mean, (2) increases the effective number of degrees of freedom, and (3) improves the statistical confidence in the results (Muller 1993; Riley 2003), and enabled us not to lose information from the residual oscillatory activity. Each SP from the sequence of SPs for one gradiometer was averaged with the corresponding SP from the sequence of SPs for another gradiometer. These averaged SPs formed the multitude of objects for further classification procedure.

The composition of brain oscillations (in terms of EEG SPs) for each type of post-stimulus interval sequence (AV(S), AV(C), AV(I), A(S), A(D), V(S), V(D)) was estimated using a probability-classification analysis of the short-term MEG SPs (SPclass tool, see Fingelkurts et al. 2006c). Considering that a detailed description of this analysis has already been published elsewhere (Fingelkurts et al. 2003a), we are highlighting here only the most important steps. In short, this analysis was undertaken in two stages (Fig. 1).

During the first stage, sequential single MEG SPs were adaptively classified in each MEG location for each subject separately using a set of standard SPs (which were generated from the data itself) (Fingelkurts et al. 2003a). As a result of the adaptive classification technique, each current SP was labeled according to the index of the class to which it belonged. Hence, each EEG signal was reduced to a sequence of individually classified SPs.

Since the length of the signal is not equal between conditions (oddball paradigm), data processing was performed separately for each 1-min portion of the signal, thus avoiding any possible bias towards representing SPs for the standard stimuli. In such way, at the second stage, the probability classification profile (PCP) of SPs for each location of 1-min MEG in each subject and for each type of post-stimulus interval was calculated (Fig. 1; see also Fig. 2 B). This index was calculated by taking the relative number of cases of an SP type as a percentage of the total amount of all SPs within each MEG location - presented as the histogram of the relative presence of each SP type (Fingelkurts et al. 2003a). It has been demonstrated in previous studies that PCP is highly stable over time (Fingelkurts et al. 2006a) and provides an adequate and detailed description of electromagnetic brain activity (Kaplan et al. 1999; Fingelkurts et al. 2003a,b; 2006b).

Analyzing the coefficient of variability (CV = standard deviation / Mean) of PCPs within each subject separately for each type of post-stimulus interval and MEG location demonstrated a very high stability of PCPs (CV = 0.0-0.6 for different MEG locations). Since the inter-individual variability of PCPs was also very low (r = 0.70-0.98 for different subject’s pairs), PCPs for each type of post-stimulus interval sequence were averaged across 213 (for AV(S)), 9 (for AV(C)), 9 (for AV(I)), 64 (for A(S)), 5 (for A(D)), 85 (for V(S)), and 6 (for V(D)) 1-min MEG signals separately for each location. It was expected that these PCPs would make it possible to illustrate in detail (in SP description) the composition of brain oscillations for the totality of the post-stimulus intervals, categorized according to the preceding stimulus. We studied the behavior of each type of spectral patterns separately and did not make any conclusions *per se* about any differences between PCPs.

In order to reveal any statistically significant differences in the relative presence of each SP type in PCPs between certain conditions, the Wilcoxon test was applied separately for each type of SPs present in the PCPs. Statistical significance was assumed where $P < 0.05$ (only statistically significant values are displayed). Since we intended to assess each variable in its own right, no Bonferroni correction was applied (for the problems associated with Bonferroni adjustments, see Perneger 1998). The decision not to make adjustments for multiple comparisons will lead to fewer errors of interpretation when the data under evaluation are not random numbers but actual observations of nature (Rothman 1990). The consistency and reproducibility of the results in the present study (results were reproduced for each subject, see above) demonstrate that the results obtained cannot be occasional because “…by definition chance findings do not replicate” (Duffy et al. 1994, p. XI). At the same time, in the case when pair test was applied between groups AV(S), AV(C), AV(I), A(S), A(D), V(S), and V(D), a Bonferroni correction was made. $P_{corrected}$ is the value required to keep the number of false positives at $P = 5\%$ (only statistically significant P-values are displayed).

3. Results

3.1. General characteristics of the probability classification profiles

Each PCP portrays the brain oscillatory patterns (in SP description), induced by particular stimuli in a particular MEG location. It should be noted that short-term power SPs were calculated for a range of 3-30 Hz (see Methods section), so the range of the brain oscillatory patterns was within the limits of this frequency.

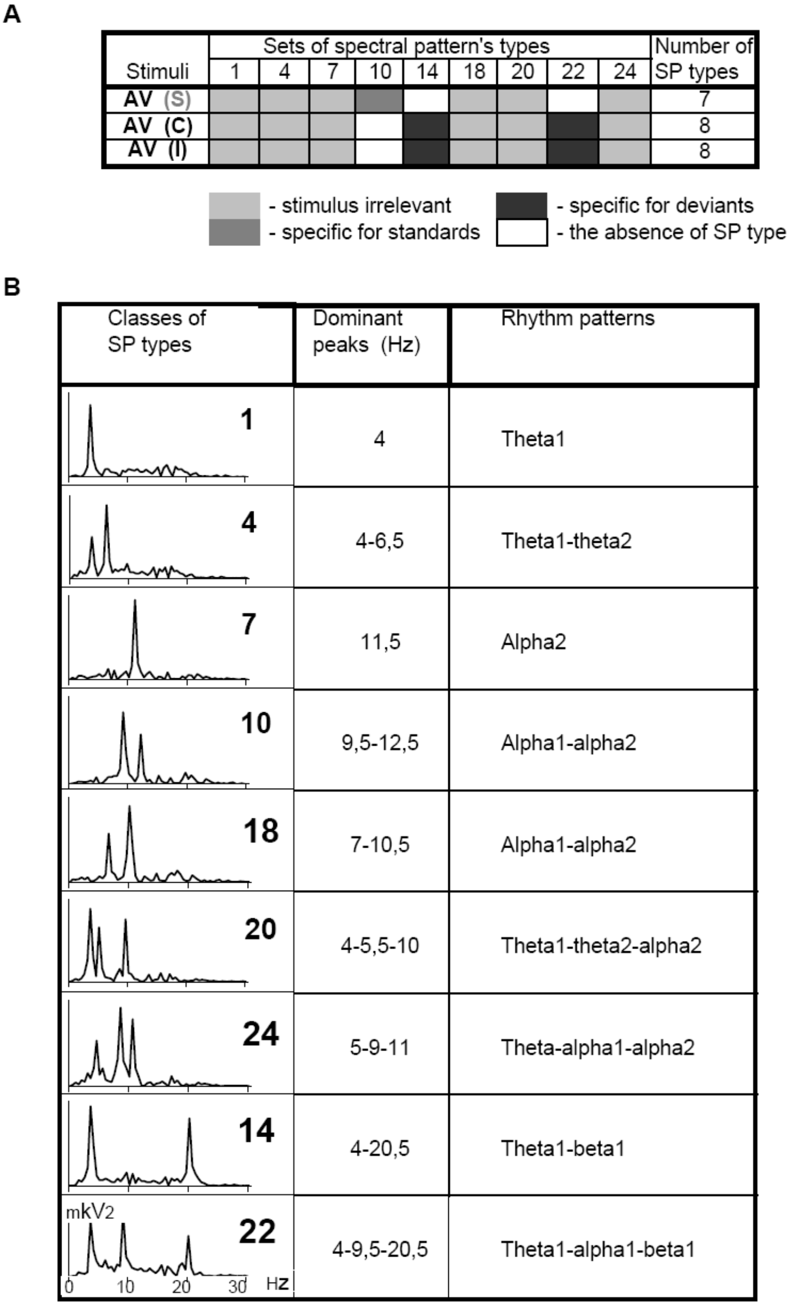
Different percentages for different SP types, mapped in PCPs, characterize the probability of particular brain oscillations occurring and/or their representation in the totality of the post-stimulus intervals, categorized with respect to the preceding stimulus. All SP types for each PCP were divided into three groups: I group - consists of SP types, each representing more than 15% of the totality of the post-stimulus intervals; II group - includes SP types, each representing between 5 and 15% of the totality of the post-stimulus intervals; and III group - comprises SP types each representing between 1 and 5 % of the totality of the post-stimulus intervals (Fig. 2, B). Only SP types from groups I and II will be included in further analyses, as the SP types which all together characterized the majority of the totality of the post-stimulus intervals.

Analysis of the PCPs demonstrated that brain oscillatory patterns induced by the stimuli could be categorized into a limited number of SP types: from 6 to 15 for different MEG locations. Different SP types had different importance for the brain when presented with the same and different stimuli. Thus, several SP types (up to 5) were the most probable when a particular type of the stimuli (Fig. 2, B) was presented. The most probable SP types are interpreted as the most “preferred oscillations” of the brain when responding to stimulus. They were estimated as the main/dominant SP-peaks in the PCP (Fingelkurts et al. 2003b, see also Barry et al. 2006).

3.2. The most probable spectral patterns for different types of stimulation

3.2.1. Audio-visual modality

Figure 3 illustrates the most probable SP types for the three types of audio-visual stimuli (AV(S), AV(C), and AV(I)). It can be seen that the sets of SP types for these stimuli did not vary significantly: the majority of the SP types were the same for all stimulus types (Fig. 3 A). However, the standard stimuli induced SP10 (with the main peaks at 9.5 and 12.5 Hz) in addition to the other SP types which were common across all three stimulus types. Whereas, the deviant (congruent and incongruent) stimuli induced SP14 (4 and 20.5 Hz) and SP22 (4, 9.5 and 20.5 Hz).



A

Stimuli	Sets of spectral pattern's types									Number of SP types
	1	4	7	10	14	18	20	22	24	
AV (S)	░	░	░	▒	□	░	░	□	░	7
AV (C)	░	░	░	□	■	░	░	■	░	8
AV (I)	░	░	░	□	■	░	░	■	░	8

░ - stimulus irrelevant
▒ - specific for standards
■ - specific for deviants
□ - the absence of SP type

B

Classes of SP types	Dominant peaks (Hz)	Rhythm patterns
1	4	Theta1
4	4-6,5	Theta1-theta2
7	11,5	Alpha2
10	9,5-12,5	Alpha1-alpha2
18	7-10,5	Alpha1-alpha2
20	4-5,5-10	Theta1-theta2-alpha2
24	5-9-11	Theta-alpha1-alpha2
14	4-20,5	Theta1-beta1
22	4-9,5-20,5	Theta1-alpha1-beta1

Fig. 3. (A) Sets of spectral pattern (SP) types each of which characterized more than 15% and/or 5-15% of the post-stimulus intervals for audio-visual standard (AV(S)), audio-visual deviant-congruent (AV(C)) and audio-visual deviant-incongruent (AV(I)) stimuli and were present in the PCPs for not less than 6 MEG locations (~ 10% of all MEG locations). (B) Represents main frequency peaks and rhythm’s patterns for corresponding SP types presented at (A).

Although the majority of the SP types were the same for all three types of stimuli, the probability of their occurring in the totality of post-stimulus intervals for different stimuli differed (Table 1). The probability of the occurrence of SP1 (4 Hz), SP20 (4, 5.5 and 10 Hz), SP14 (4 and 20.5 Hz) and SP22 (4, 9.5 and 20.5 Hz) was significantly higher for the deviant stimuli (both congruent and incongruent) when compared with standard stimuli (*P*<0.007-*P*<0.00001 for different MEG locations). At the same time, the probability of SP18 (7 and 10.5 Hz) and SP24 (5, 9 and 11 Hz) occurring was significantly lower for the deviant stimuli (both congruent and incongruent) when compared with standard stimuli (*P*<0.007-*P*<0.0001 for different MEG locations) (Table 1). This effect was observed for 50-85% (for different SP types) of MEG locations (Fig. 4).

Table 1. Spectral patterns (SP) which showed changes in the probability index for audio-visual deviant (D) (both congruent and incongruent) stimulation when compared with audio-visual standard (S) stimulation. Data averaged across all subjects and represent the effect in 50-85% (for different SP types) of MEG locations.

SP type	Main peaks (Hz)	Rhythm patterns	S *vs* D	Significance
1	4	theta	<	*p*<0.007-*p*<0.00001
20	4-5.5-10	theta1-theta2–alpha	<	*p*<0.007-*p*<0.00001
14	4-20.5	theta–beta	<	*p*<0.007-*p*<0.00001
22	4-9.5-20.5	theta–alpha–beta	<	*p*<0.007-*p*<0.00001

a

18	7-10.5	alpha1-alpha2	>	*p*<0.007-*p*<0.0001
24	5-9-11	theta–alpha1-alpha2	>	*p*<0.007-*p*<0.0001

b

Although this effect was common to both AV(C) and AV(I) stimuli, the probability index wasn’t the same for these two deviant stimulus types (not shown). The probability of each of the SP types (shown in the Table 1) occurring was significantly higher for the deviant-incongruent stimuli compared to a deviant-congruent one relative to the standard stimuli: AV(S) < AV(C) < AV(I) (*P*<0.007-*P*<0.00001 for different SP types and MEG locations). For those SP types present during the deviant stimuli in which the probability index was significantly lower than for the standard stimuli, this index was significantly lower for the deviant-congruent stimuli than for deviant-incongruent: AV(S) > AV(I) > AV(C) (*P*<0.007-*P*<0.0001 for different SP types and

MEG locations). These results suggested that conflictual auditory and visual information of AV(I) stimuli elicited specific PCP (in terms of probability index) when compared with AV(C) stimuli.

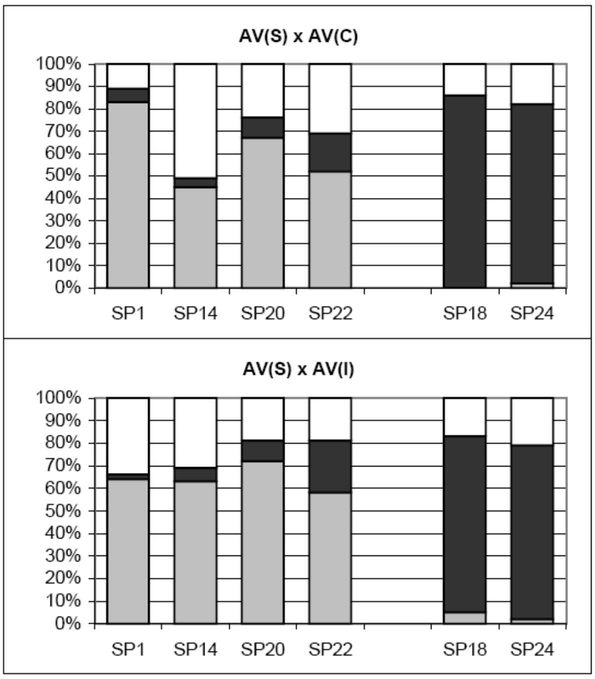


Fig. 4. The number of MEG locations (in percents, 100% – 64 locations) which revealed changes in the probability index for particular spectral pattern (SP) types for audio-visual congruent (AV(C)) and audio-visual incongruent (AV(I)) stimuli when compared with audio-visual standard (AV(S)) stimuli.
White colour – the number of locations, where there were no changes; gray colour – the number of locations, where the probability index for particular SP types during AV(C) and AV(I) stimulation was larger when compared with AV(S) stimulation; black colour – the number of locations, where the probability index for particular SP types during AV(C) and AV(I) stimulation was smaller when compared with AV(S) stimulation.

3.2.2. Auditory and Visual Modalities

For the auditory and visual modalities, only standard and deviant stimuli were presented to the subjects. An entire comparison was made of SP sets for both modalities and all types of stimulation. In these modalities, the presence of SP1 (4 Hz), SP4 (4 and 6.5 Hz), SP18 (7 and 10.5 Hz), and SP24 (5, 9 and 11 Hz) did not depend on the modality or stimulus type. SP7 (11.5 Hz) and SP20 (4, 5.5 and 10 Hz) were common to both of the modalities and to the majority of the stimuli, however, SP7 was not observed for V(D) stimuli and SP20 was not present during A(S) stimuli. In addition, SP10 (9.5 and 12.5 Hz) was the most probable SP only during the presentation of the auditory deviant stimuli. SP14 (4 and 20.5 Hz) was present in auditory and visual deviant conditions while SP22 (4, 9.5 and 20.5 Hz) was only present in A(D) condition. At the same time, there were a few SPs which were specific to each of the modalities. Thus, SP2 (5.5 Hz) was presented only during the presentation of V(D) stimuli and SP12 (28.5 Hz) and SP13 (4 and 9 Hz) were found only for A(D) stimuli.

3.2.3. Comparative analysis of Audio-Visual, Auditory and Visual Modalities

A comparative analysis of the probability of SP types occurring was undertaken separately for each MEG location for the following combinations – modality(stimulus type): A(S) x A(D), V(S) x V(D), AV(C) x A(D), AV(C) x V(D), A(D) x V(D), AV(I) x A(D), AV(I) x V(D), AV(S) x A(S), AV(S) x V(S), A(S) x V(S). [Comparisons AV(S) x AV(C), AV(S) x AV(I) and AV(C) x AV(I) were considered above (see section Audio-visual modality)]. The results of this analysis showed that the probability of SP types with mostly theta components occurring was the greatest for V(D) stimuli, lower for AV(C), AV(I) and V(S) stimuli, followed by AV(S) and A(D), and the lowest for A(S) stimuli ($P_{corrected}$<0.007-$P_{corrected}$<0.00001 for different MEG locations) (Table 2). SP types with mostly alpha components showed the opposite results in terms of the probability index ($P_{corrected}$<0.007-$P_{corrected}$<0.0001 for different MEG locations) (Table 2). Comparisons made revealed that the visual-modality had a significantly higher probability index for SP types with mostly theta components and a significantly lower probability index for SP types with alpha components when compared with the auditory-modality ($P_{corrected}$<0.007-$P_{corrected}$<0.0001 for different MEG locations). These effects were shown for 71% of the MEG locations.

Table 2
Comparisons of the probability index separately for spectral patterns (SP) with mostly theta and alpha components for different types of stimulation. Data averaged across all subjects and represent the effect in the majority of MEG locations.
Bonferroni correction was used.

SPs with mostly theta components	$A(S) \underset{p<0.007\text{-}0.00001}{<} \left\{\begin{matrix} AV(S) \\ A(D) \end{matrix}\right\} \underset{p<0.007\text{-}0.00001}{<} \left\{\begin{matrix} AV(C) < AV(I) \\ V(S) \\ p<0.007\text{-}0.00001 \end{matrix}\right\} \underset{p<0.007\text{-}0.00001}{<} V(D)$
SPs with mostly alpha components	$A(S) \underset{p<0.007\text{-}0.0001}{>} \left\{\begin{matrix} AV(S) \\ A(D) \end{matrix}\right\} \underset{p<0.007\text{-}0.0001}{>} \left\{\begin{matrix} AV(I) > AV(C) \\ V(S) \\ p<0.007\text{-}0.0001 \end{matrix}\right\} \underset{p<0.007\text{-}0.0001}{>} V(D)$

AV(S) - audio-visual standard stimuli; V(S) – visual standard stimuli; A(S) – acoustic standard stimuli; AV(C) - audio-visual deviant-congruent stimuli; AV(I) - audio-visual deviant-incongruent stimuli; V(D) – visual deviant stimuli; A(D) – acoustic deviant stimuli.

Moreover, six SP types (1, 4, 7, 14, 20 and 24) from each of the stimulus type SP sets revealed reciprocal behavior in the probability index for auditory and visual modalities when compared to any other modality(stimulus type). These SPs partially determined the opposite behavior of the probability index for auditory and visual modalities.

Conventional ‘energetic’ estimation (mean spectral power) of the stimuli-related MEG SPs, averaged across all inter-stimulus intervals for each stimulus type separately, revealed similar effects in some MEG locations (central and temporal areas) but they did not reach statistical significance due to the very high variability of the SPs (see Fig. 5).

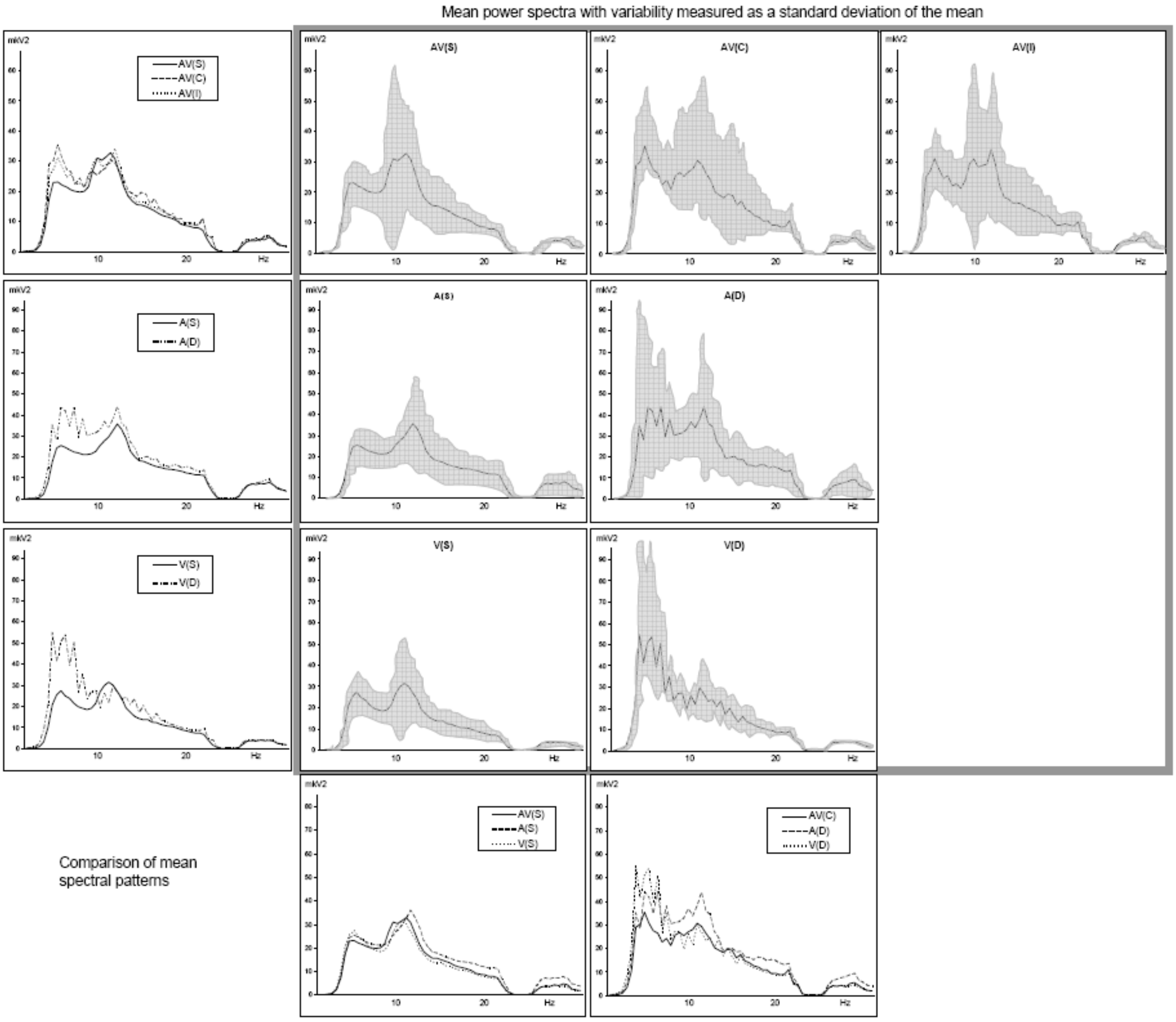


Fig. 5. Example of conventional mean power spectra of the stimuli-related spectral patterns averaged across all inter-stimulus intervals for each stimulus type separately for MEG 2423/2422 (T6) (probability-classification profiles for this location are presented at Fig.2.B).
AV(S) – audio-visual standard stimuli; AV(C) – audio-visual deviant-congruent stimuli; AV(I) – audio-visual deviant-incongruent stimuli; A(S) – auditory standard stimuli; A(D) auditory deviant stimuli; V(S) – visual standard stimuli; V(D) – visual deviant stimuli. The gray areas encompass variability of the spectral patterns measured as a standard deviation of the mean.

4. Discussion

4.1. Methodological aspects

In attempting to answer the question of how an integrated sensory percept is represented throughout the cortex and which brain oscillations are involved in its maintenance, we focused on brain oscillations within the 3-30 Hz range. The usage of sliding spectral analysis on the basis of Short-Time FFT (see parameters at Methods section) enabled us to achieve a reasonable time-frequency resolution along all frequencies studied and to obtain an exact view of the contribution of frequency components to probability classification profiles. Time-frequency representation analysis (TFR) cannot be used for the purpose of the present study since no algorithm for classification of the individual TFR-maps exists at the moment, and the creation of such an algorithm would not be an inconsiderable undertaking. By means of a probability-classification technique, based on estimating the relative presence of SP types in the totality of post-stimulus intervals, we assessed ‘internal’ events as the probability of brain oscillatory patterns induced by AV, V and A stimuli. Supposing (a) that long-lasting brain oscillations induced by the same stimulus events are systematic and (b) that unspecific neurophysiologic processes that take place in the same time window are random during each trial, we believe that the post-stimulus intervals selected for the analysis most likely reflect a maintained final, multimodally determined conscious percept (see Introduction). Being a complex neurophysiological state, maintenance of percept involves some cognitive aspects of stimulus processing, such as attention, encoding, and sensory short-term memory (see reviews Mesulam 1998; Näätänen and Winkler 1999; see also Basar et al. 2004).

In the present pilot study, all subjects demonstrated very similar spectral pattern variability (in terms of PCPs, r = 0.70-0.98 for different subject pairs) for a particular experimental condition (V, A or AV stimulation). In another words, the composition and percent ratio of SP types in PCPs were very similar among the subjects. This can be explained by the fact that the task itself and subject’s focused attention by their very nature reduced any variability in the functional states of the brain, and thereby in MEG characteristics (McEvoy et al. 2000).

4.2. The probability of occurrence of spectral patterns induced by uni- and bimodal stimulation

The PCPs obtained for all MEG locations during A, V, and AV stimulation revealed the complicated structure of brain magnetic oscillatory responses at different frequencies. Thus, several peaks in PCPs (see Fig. 2.B) reflect the fact that oscillatory MEG responses from different frequency bands (independent of modality) are generated after the external stimulation. This finding is consistent with previous publications (Stampfer and Basar 1985; Basar 1999; Karakas et al. 2000). Brain oscillatory pattern types irrespective of modality (AV, A, V) and stimulus type (S, D, C, I) were found in the present study. At the same time, the probability index of these SP types differed significantly for different modalities and stimulus types reflecting the variable but not random nature of brain oscillations (Fingelkurts et al. 2003b). The probability index in the present study reflects the chance of particular brain oscillations being elicited in the totality of post-stimulus intervals categorized with respect to the preceding stimulus. Additionally, specific SP types for deviant (all modalities) and standard stimuli (AV modality) were found.

Although the subjective/conscious mental percepts for AV(C) and AV(I) stimuli were the same, their neuronal representations differed significantly. Thus, the probability of SPs with mostly theta and beta1 components occurring was higher when the AV deviant (especially incongruent – the McGurk-type) stimuli were presented when compared with the AV standard stimuli: AV(S) < AV(C) < AV(I). Conversely, the probability of SPs with mostly alpha components occurring was higher when AV standard stimuli were presented when compared with the AV deviant stimuli (especially congruent): AV(S) > AV(I) > AV(C) (Table 1). It seems that congruent auditory information enhances alpha blockage elicited by the visual information. The same difference between standard and deviant stimuli was found for both A and V modalities. However, it was more pronounced during the presentation of V stimuli when compared with the stimulation in the AV and A modalities (Table 2). Such reciprocal relationships between theta and alpha rhythms probably reflect the situation when standard and deviant stimuli triggered two different oscillatory generator systems (Klimesch 1996).

It seems that V and A modalities are coordinated by different yet interrelated oscillatory networks: theta for auditory, and alpha for visual modalities (Table 2). These results support the fact of the existence of alpha blockade during visual stimulation. AV modality was driven by the combination of these two brain oscillations which demonstrated reciprocal relations with the domination of oscillatory behavior, typical for visual modality (Table 2). Thus, the probability index during audio-visual modality was medium when compared with only auditory and only visual stimulation: A<AV<V for SPs containing mostly theta components and A>AV>V for SPs containing mostly alpha components. These results are somewhat different from the findings in the work of Yordanova et al. (2002). They demonstrated a strong predominance of the relative theta power in each modality (A, V and AV). This difference from the results of the present study can be explained by the fact that Yordanova and colleagues used simple stimuli (tone bursts and rectangular light-stimuli) in a passive condition. Whereas, in the present study we used complex natural stimuli (a talking face) and a task which involved attention (see Material and methods). As a consequence, the overall picture of the frequency components involved and their interaction is more complicated in the present study.

Additionally, stimulation with the deviant stimuli (all modalities except visual) revealed a beta1 (around 20 Hz) oscillatory contribution. Note that this oscillation never operated independently: it was always accompanied by theta and/or theta-alpha oscillations. Stimulation with A(D) stimuli was characterized by beta2 (around 28-29 Hz) oscillations mainly in the frontal locations.

The majority of the most probable SP types contained two and/or three main frequency peaks. Thus, frequency components were generally non-independent. This supports the idea of the existence of dependent oscillatory generators (Bullock 1997).

Below is a discussion of possible functions of the brain oscillations induced by uni- and bimodal stimulation.

4.3. Functional interpretation of frequency components within spectral patterns induced by uni- and bimodal stimulation

4.3.1. Alpha oscillations

It is well known that alpha is more than just a spontaneous rhythm (for functional correlates see, Basar et al. 2001a) and may be a prototype of the ‘building blocks’ (Lehmann 1989) which accompany psycho-physiological events. A decrease in the probability index for SPs containing mostly alpha components for deviant stimuli (all modalities) when compared with standard stimuli (all modalities) may indicate that perceptual processes and stimulus encoding have been activated, in the same way as Klimesch (1999) demonstrated for alpha rhythm desynchronization. It appears that the habituation processes (decreased reactivity) due to refractoriness of the neuronal circuits, responding to the frequent standard stimulus, was reflected in the large probability index for SPs with mostly alpha components. Moreover, SP10 containing frequencies from upper alpha (main peaks at 9.5 and 12.5 Hz) was found only for the standard stimuli (all modalities). Maintenance of infrequent deviant stimulus resulted in the activation of fresh afferent neuronal elements (Näätänen and Winkler 1999). This may have been reflected in the appearance of “new” brain oscillations: mostly theta and/or theta-beta (Table 1.a). The fact that the probability index of SPs containing components from low and upper alpha were affected (decreased) when deviant stimuli were presented may reflect the allocation of attention (low alpha) and task-specific cortical activation (upper alpha) (Pfurtscheller and Klimesch 1990) as well as the activation of two neuronal networks: the thalamo-cortical and cortical respectively (Lopes da Silva et al. 1980).

It has been shown earlier that alpha activity decreases during selective attention (Steriade et al. 1990; Lopes da Silva 1991). The fact that the probability of alpha oscillations being elicited in the present study during the audio-visual stimulation was less than during the auditory and more than during the visual stimulations (A>AV>V) may suggest that the maintenance of audio-visual stimuli involves more focused attention than the maintenance of auditory stimuli, but still less than the maintenance of visual stimuli.

4.3.2. Theta oscillations

Activity in the theta band (in contrast to the alpha band) may be responsible for the encoding of new information and associated with short-term memory processes (Doppelmayr et al. 1998;

Klimesch 1999; for the review see, Schacter 1977) which are essential for the maintenance of the information. Reverting to the data outlined above on the probability index for SPs with mostly theta components this index was higher during the presentation of deviant stimuli (all modalities) when compared with standard stimuli (Tables 1.a, 2). All these SPs contained 4 Hz, and some of them had additional peaks at 5.5 Hz or 6.5 Hz. This suggests that a greater probability index for SPs with mostly theta components during deviants when compared with standards may also indicate that sensory short-term memory processes are active during the maintenance of the deviant percepts. This implies a process which detects the deviation with reference to the traces encoding the most recent history of the previous stimulation. Perhaps, data collected in the present study reflect a change-detection process based on the information stored in these traces.

There is also the possibility that theta declines as tasks become more familiar/easier (Gundel and Wilson 1992). This may explain the fact that the probability index of SPs with mostly theta components for AV(S), A(S), V(S) (familiar events) was less than for AV(C, I), A(D), V(D) (novel events) (Table 2). We assume that attention may contribute to this difference. According to recent theories of automaticity, performance of a new task is based on a general algorithm for solving the problems of the task and attends to the various steps of this algorithm, whereas automatic performance is based on single-step, direct-access retrieval of past solutions from memory, thus attending to the solutions that memory presents (for a review, see Logan 1992). AV-deviant stimuli consist of auditory and visual components and it has been shown that attention plays a different role in the processing of acoustic and visual information (Treisman and Sato 1990; Näätänen and Winkler 1999). An involuntary diversion of the subject’s attention away from the primary task, apparently caused by the neuronal-mismatch response elicited by the acoustic deviant stimuli, has been clearly demonstrated (Näätänen and Winkler 1999). Feature integration forming unitary auditory perceptual events most likely occurs with the participation of an involuntary shifting of the subject’s attention. In other words, attention can be attracted automatically.

Perceptual stimulus formation in visual-modality has an important difference from auditory-modality (Treisman and Sato 1990). According to Treisman’s theory, separable visual features are registered independently and in parallel, i.e. automatically, but focused attention is needed to combine accurately the features of a complex stimulus. Further support for this theory can be found in a recent review (Kanwisher and Wojciulik 2000) which emphasizes that the processing of visual information depends critically on attention.

Obviously, these peculiarities in the processing of acoustic and visual information should merge during an audio-visual stimulation, which was notably revealed in the present study (Table 2; see SPs containing mostly theta components). The probability index of SPs with mostly theta components was greatest during the presentation of stimuli in visual-modality, the smallest in auditory-modality and moderate in audio-visual-modality (at the same time, AV(C) < AV(I)). This data may be associated with the fact that visual information has a significant and dominant influence on audio-visual information processing (Sams et al. 1991; Rosenblum and Saldana 1996) and its maintenance. With attention as a necessary component for the accurate integration of stimulus features (for a review, see Kanwisher and Wojciulik 2000), visual modality involuntary attracts a subject's attention making the visual component in audio-visual stimulus dominant. This idea is supported by a recent study (Tiippana et al. 2004) in which the authors studied the role of visual attention in audiovisual speech perception by measuring the McGurk effect. They demonstrated that visual attention strongly modulates audio-visual speech perception and suggested that attention is required to combine visual and auditory speech features: "If attention is disrupted, audiovisual integration is less efficient even though the processing of unisensory features remains intact" (Tiippana et al. 2004). The dominance of visual information may be explained also by the view that visual information is considered instructive in the formation of the sensory representations for the other modalities (Knudsen and Brainard 1995).

The fact that the probability index for SPs containing mostly theta components was larger for deviant-incongruent (McGurk-type) stimuli when compared with deviant-congruent stimuli suggests that the maintenance of incongruent stimulus (more difficult task) requires the allocation of more attentional resources than the congruent one (more simple task) (Gundel and Wilson 1992).

One counterargument to our interpretation can be derived from the recent study by Colin et al. (2002) which measured the mismatch negativity (MMN) attributed to the auditory illusion produced by McGurk stimuli. They suggested that audio-visual integration occurs pre-attentively. However, it is unclear whether the MMN correlates observed by Colin et al. were produced by the consequences of audio-visual integration or directly by the consequences of a discrepancy between the auditory and the visual stimuli. Soto-Faraco et al. (2004) have reported on a series of experiments designed to assess directly the automatic nature of integration during the McGurk effect. Their results support the hypothesis that the McGurk illusion occurred automatically. However, according to recent theories of automaticity, automatic does not mean attention-free (see above, and Logan 1992). Thus, a recent review of available literature (Calvert and Thesen 2004) suggests that involuntary shifts in attention may arise as a consequence of multisensory integration and that some of the attentional mechanisms may influence the integration and maintenance of multisensory cues. This view supports our data and interpretation. The “fresh” study of Alsius et al. (2005) in which the authors demonstrated that in contrast to the assumption that multimodal speech integration is automatic, these multisensory binding processes are subject to attentional demands, is also consistent with our discussion.

4.3.3. Beta oscillations

SP14 (main peaks at 4 and 20.5 Hz) (regardless of modality) and SP22 (4, 9.5 and 20.5 Hz) (except visual-modality, most likely because of alpha component) which were specific for the deviant stimulus were found in the present study. Both of these SPs were characterized by a beta component. It is interesting to note that the probability index of these SPs was the largest for deviant-incongruent stimuli (McGurk-type). It is known that beta rhythm synchronizes after finger, arm, foot (Pfurtscheller et al. 1996), mouth (Salmelin et al. 1995) and face (Arroyo et al. 1993) movement, as well as after imagined movement (Neuper and Pfurtscheller 1999). Oscillations ~20 Hz are usually considered harmonic with the mu rhythm, acting in the same way as ~10 Hz frequency for the mu rhythm, and being strictly localized (Hari and Salenius 1999). However, our data on SPs containing ~20 Hz revealed differences from these mu rhythm properties at 20 Hz.

Thus, it was shown in the present study that the probability index for SPs containing 20.5 Hz always ‘behaved’ in the opposite way to the probability index of SPs containing 10 Hz (and/or other alpha components). The frequency components of SPs 14 and 22 reveal that they represent separate rhythms, (albeit highly functionally interconnected ones, i.e. always occurring together), rather than harmonically related frequencies. These SPs were found for the majority of MEG locations (up to 63% for different SPs and stimulus types, Fig. 4). Note also that SPs with a 20.5

Hz component weren’t detected for standard stimuli (standard stimuli except for the auditory-modality contain lips movement). Hence, all these peculiarities of the 20.5 Hz component for SPs 14 and 22 suggest that the 20.5 Hz component found in the present study most likely represents a beta rhythm which is separate from the mu rhythm and that this beta rhythm is particularly important for the maintenance of a conscious audio-visual percept. This finding is consistent with the work of Krause et al. (2001) where the authors demonstrated that brain oscillations ~ 20 Hz (beta rhythm) responded to audio-visual speech information.

Altogether our findings suggest that the maintenance of the audio-visual integrative percept as well as the maintenance of the percepts in auditory and visual modalities is accompanied by an interplay of several brain oscillations: alpha, theta and beta rhythms which have different functional operations (see also Basar et al. 2001a).

4.4. Distributed oscillatory networks of brain oscillatory patterns induced by uni- and bimodal stimulation

The main effect described above (Table 1, 2) was detected in the majority (up to 85%) of MEG locations. This may reflect the involvement of distributed neuronal networks in the maintenance of the auditory, visual and audio-visual percepts of talking face (Fig. 4). This finding is in line with those of Raij (1999) and Sakowitz et al. (2001), according to whom cortical activation following audio-visual stimuli involves the activation of multiple cortical regions. Also, it was shown that even simple visual stimuli activate a large number of distinct occipital, temporal, and parietal cortical areas (Felleman and Van Essen 1991). The existence of distributed alpha (for a review, see Basar et al. 2001a) and theta (Basar et al. 2001b) oscillatory networks has been well documented. Since the main effects described in the present paper have been observed in the majority of MEG locations and in connection with the maintenance of different types of stimuli, it would not be plausible to attribute our findings to one specific system. The distributed nature of the effects observed in the present study supports the idea that a cognitive function such as the maintenance of the integrative sensory percept is based on the extraordinary interconnectivity of the whole cortex (Tononi et al. 1998; Gazzaley et al. 2004; Fingelkurts and Fingelkurts, 2005) and is organized into parallel processing streams with complementary properties (Grossberg 2000). Probably, these peculiarities provide conditions for the generation of a metastable multisensory scene (Fingelkurts et al. 2003c; for the review on brain metastability, see Fingelkurts and Fingelkurts 2004).

Perhaps distributed theta, alpha and beta oscillatory systems observed in the present study act as resonant communication networks through large populations of neurons (for a review, see Basar et al. 2001a). These resonant communication networks (distributed oscillatory systems) may provide a general communication framework parallel to the morphology of distributed sensory networks (for a review, see Basar et al. 2001a). In other words, the temporal properties of cortical inputs could determine which resonant modes were activated, effectively specifying the spatial distribution of correlated neural activity throughout the brain (Fingelkurts et al. 2003c). The existence of distributed theta, alpha and beta oscillatory systems involved in maintaining the auditory, visual and audio-visual percepts which were found in the present study is consistent with the findings on functional synchronization between MEG locations analyzed from the same data earlier (Fingelkurts et al. 2003c). The authors demonstrated that audio-visual synthesis takes place within a distributed and dynamic cortical network with emergent and metastable properties, thus supporting a model of crossmodal integration suggested by Calvert (2001) and operational model suggested by Fingelkurts and Fingelkurts (2005, 2006).

Due to the small number of subjects, the findings of this pilot study are preliminary and should be validated with a larger number of subjects. However, consistency and reproducibility of the results (the results were reproduced for each subject) demonstrate the reliability of the presented findings. Further experiments will be necessary to (a) distinguish task-specific processes (attention, encoding, and sensory short-term memory) which contribute to maintenance state and (b) determine both the sources of examined activity as well as its specific spatial properties.

5. Conclusions

To summarize, the probability-classification technique used in the present exploratory study illuminates the complexity of brain oscillatory processes underlying the maintenance of the auditory, visual and integrated audio-visual percepts. It was demonstrated that for all percept modalities spectral pattern types were of varying significance: several spectral pattern types were more probable than others. The probability index was dependent on the type and the modality of the sensory percept.

The data obtained in the present study suggests the existence of distributed oscillatory networks whilst auditory, visual and audio-visual percepts are maintained. Oscillations around 20 Hz were characteristic for an integrated audio-visual percept. This indicates that the maintenance of a multisensory percept is not mediated by a unitary physiological mechanism but results from multiple neuronal operations and cannot be attributed to a separate individual brain area.

Identifying the actual composition of brain oscillations has allowed us (a) to distinguish two subjectively and consciously identical mental percepts, and (b) to characterize the types of brain functions involved in maintenance of the multisensory percept.

Maintenance of the integrated audio-visual percept was accompanied by a unique composition of distributed brain oscillations, typical for auditory and visual modalities; and the contribution of brain oscillations characteristic for visual modality was dominant.

Acknowledgements

The authors thank Riikka Möttönen for the MEG registration. MEG recordings were made in the Brain Research Unit at the Low Temperature Laboratory at the Helsinki University of Technology. The authors wish to thank Carlos Neves, computer science specialist, for the software development and technical support. Special thanks to Simon Johnson for skilful language editing. AAFs were supported by Research Fellowship from CIMO, Finland. Preparation of this article was supported by BM-SCIENCE. The authors state that experiment of this study complied with the current laws of Finland.

References

Alsius A, Navarra J, Campbell R, Soto-Faraco S (2005) Audiovisual integration of speech falters under high attention demands. Curr Biol 15(9): 839-843

Arroyo S, Lesser RP, Gordon B, Uematus S, Jackson D, Webber R (1993) Functional significance of the mu rhythm of human cortex: an electrophysiological study with subdural electrodes. Electroencephalogr Clin Neurophysiol 87: 76-87

Basar E (1999) Brain function and oscillations. Vol. II. Integrative Brain Function. Neurophysiology and Cognitive Processes. Springer, Berlin

Basar E, Basar-Eroglu C, Karakas S, Schurmann M (2001a) Gamma, alpha, delta, and theta oscillations govern cognitive processes. Int J Psychophysiol 39: 241-248

Basar E, Schurmann M, Sakowitz O (2001b) The selectively distributed theta system: functions. Int J Psychophysiol 39: 197-212

Basar E, Schurmann M, Demiralp T, Basar-Eroglu C, Ademoglu A (2001c) Event-related oscillations are ‘real brain responses’ – wavelet analysis and new strategies. Int J Psychophysiol 39: 91-127

Basar E, Özgören M, Karakas S, Basar-Eroglu C (2004) Super-synergy in the brain: the grandmother percept is manifested by multiple oscillations. Int J Bifurcat Chaos in Applied Sciences and Engineering 14: 453-491

Barry RJ, Rushby JA, Janette L, Smith JL, Clarke AR, Rodney J, Croft RJ (2006) Dynamics of narrow-band EEG phase effects in the passive auditory oddball task. Eur J Neurosci 24: 291

Bullock TH (1997) Signals and signs in the nervous system: The dynamic anatomy of electrical activity. Proc Natl Acad Sci U S A 94: 1-6

Calvert GA (2001) Crossmodal processing in the human brain: Insights from functional neuroimaging studies. Cereb Cortex 11: 1110-1123

Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. J Physiol (Paris) 98: 191–205

Colin C, Radeau M, Soquet A, Demolin D, Colin F, Deltenre P (2002) Mismatch negativity evoked by the McGurk–MacDonald effect: a phonetic representation within short-term memory. Clin Neurophysiol 113: 495–506

Courtney SM, Ungerleider LG, Keil K, Haxby JV (1997) Transient and sustained activity in a distributed neural system for human working memory. Nature 386: 608-611

Doppelmayr M, Klimesch W, Schwaiger J, Auinger P, Winkler T (1998) Theta synchronization in the human EEG and episodic retrieval. Neurosci Lett 257: 41-44

Driver J, Spence C (2000) Multisensory perception: beyond modularity and convergence. Curr Biol 10: R731-R735

Duffy F, Hughes JR, Miranda F, Bernard P, Cook P (1994) Status of quantitative EEG (QEEG) in clinical practice. Clin Electroencephalogr 25: VI-XXII

Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. Nature reviews. Neuroscience 2: 704-716

Ernst MO, Bulthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8: 162-169

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cerebl Cortex 1: 1-47

Fingelkurts AlA, Fingelkurts AnA, Krause CM, Sams M (2002) Probability interrelations between pre-/post-stimulus intervals and ERD/ERS during a memory task. Clin Neurophysiol 113: 826-843

Fingelkurts AlA, Fingelkurts AnA, Kaplan AYa (2003a) The regularities of the discrete nature of multi-variability of EEG spectral patterns. Int J Psychophysiol 47: 23-41

Fingelkurts AlA, Fingelkurts AnA, Krause CM, Kaplan AYa (2003b) Systematic rules underlying spectral pattern variability: Experimental results and a review of the evidences. Int J Neurosci 113: 1455-1481

Fingelkurts AnA, Fingelkurts AlA (2004) Making complexity simpler: multivariability and metastability in the brain. Int J Neurosci 114: 843–862

Fingelkurts AnA, Fingelkurts AlA (2005) Mapping of the brain operational architectonics. In: Chen FJ (ed.) Focus on brain mapping research. Nova Science Publ, pp 59-98

Fingelkurts AnA, Fingelkurts AlA (2006) Timing in cognition and EEG brain dynamics: discreteness versus continuity. Cog Proces 7: 135-162

Fingelkurts AnA, Fingelkurts AlA, Krause CM, Möttönen R, Sams M (2003c) Cortical operational synchrony during audio-visual speech integration. Brain Lang 85: 297-312

Fingelkurts AlA, Fingelkurts AnA, Kaplan AYa (2006a) Stability, reliability and consistency of the compositions of brain oscillations. Int J Psychophysiol 59(2): 116-126

Fingelkurts AlA, Fingelkurts AnA, Kaplan AYa (2006b) Interictal EEG as a physiological adaptation. Part I: Composition of brain oscillations in interictal EEG. Clin Neurophysiol 117: 208-222

Fingelkurts AA, Fingelkurts AA, Kivisaari R, Autti T, Borisov S, Puuskari V, Jokela O, Kähkönen S (2006c) Reorganization of the Composition of Brain Oscillations and Their Temporal Characteristics in Opioid Dependent Patients. Prog Neuropsychopharmacol Biol Psychiatry 30(8): 1453-1465

Friederici AD (2002) Towards a neural basis of auditory sentence processing. Trends Cogn Sci 6: 78-84

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey’s dorsolateral prefrontal cortex. J Neurophysiol 61: 331-349

Fuster JM, Alexander GE (1973) Firing changes in cells of the nucleus medialis dorsalis associated with delayed response behavior. Brain Res 61: 79-91

Gazzaley A, Rissman J, D’Esposito M (1994) Functional connectivity during working memory maintenance. Cogn Affect Behav Neurosci 4(4): 580-599

Grossberg S (2000) The complementary brain: unifying brain dynamics and modularity. Trends Cogn Sci 4:233-246

Gundel A, Wilson GF (1992) Topographical changes in the ongoing EEG related to the difficulty of mental tasks. Brain Topogr 5: 17-25

Hari R, Salenius S (1999) Rhythmical corticomotor communication. NeuroReport 10: R1-R10

Kanwisher N, Wojciulik E (2000) Visual attention: insights from brain imaging. Nat Rev Neurosci 1: 91-100

Kaplan AYa (1998) Nonstationary EEG: methodological and experimental analysis. Uspehi Physiologicheskih Nayk (Success in Physiological Sciences) 29(3): 35-55 (In Russian)

Kaplan AYa, Fingelkurts AlA, Fingelkurts AnA, Grin’ EU, Ermolaev VA (1999) Adaptive classification of dynamic spectral patterns of human EEG. Journal VND (Journal of Higher Nerve Activity) 49: 416-426 (In Russian)

Karakas S, Erzengin O, Basar E (2000) A new strategy involving multiple cognitive paradigms demonstrates that ERP components are determined by the superposition of oscillatory responses. Clin Neurophysiol 111: 1719-32

King AJ, Calvert GA (2001) Multisensory integration: perceptual grouping by eye and ear. Curr Biol 11: R322-R325

Klimesch W (1996) Memory processes, brain oscillations and EEG synchronization. Int J Psychophysiol 24: 61-100

Klimesch W (1999) Event-related band power changes and memory performance. Event-related desynchronization and related oscillatory phenomena of the brain. In: Pfurtscheller G, Lopez da Silva FH (eds). Handbook of electroencephalography and clinical neurophysiology, Vol. 6, revised ed. Elsevier, Amsterdam, pp 151-178

Knudsen EI, Brainard MS (1995) Creating a unified representation of visual and auditory space in the brain. Annu Rev Neurosci 18: 19–43

Krause CM, Möttönen R, Jensen O, Lampinen J, Sams M (2001) Brain oscillations during audiovisual speech perception – a MEG study. Paper was presented at the Human Brain Mapping Conference, June 10-14, England

Kutas M, Hillyard SA (1980) Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207: 203-205

Large M-E, Aldcroft A, Vilis T (2005) Perceptual Continuity and the Emergence of Perceptual Persistence in the Ventral Visual Pathway. J Neurophysiol 93: 3453–3462

Laskaris NA, Ioannides AA (2001) Exploratory data analysis of evoked response single trials based on minimal spanning tree. Clin Neurophysiol 112: 698-712

Lebib R, Papo D, Douiri A, de Bode S, Gillon Dowens M, Baudonnière P-M (2004) Modulations of 'late' event-related brain potentials in humans by dynamic audiovisual speech stimuli. Neurosci Lett 372(1-2): 74-79

Lehmann D (1989) From mapping to the analysis and interpretation of the EEG/EP maps. In: Maurer K (ed). Topographic brain mapping of EEG and evoked potentials. Springer, Berlin Heidelberg New York, pp 53-75

Levy WJ (1987) Effect of epoch length on power spectrum analysis of the EEG. Anesthesiology 66(4): 489-495

Libet B (1993) The neural time factor in conscious and unconscious events. Experimental and Theoretical Studies of Consciousness (CIBA Foundation Symposium 174). Wiley, Chichester, pp 123-146

Logan GD (1992) Attention and preattention in theories of automaticity. Am J Psychol 105: 317-339

Lopes da Silva FH (1991) Neural mechanisms underlying brain waves: from neural membranes to networks. Electroencephalogr Clin Neurophysiol 79: 81-93

Lopes da Silva FH, Vos JE, Mooibroek J, Van Rotterdam A (1980) Relative contributions of intracortical and thalamo-cortical processes in the generation of alpha rhythms, revealed by partial coherence analysis. Electroencephalogr Clin Neurophysiol 50: 449-456

McCollough AW, Machizawa MG, Vogel EK (2007) Electrophysiological measures of maintaining representations in visual working memory. Cortex 43: 77-94

McEvoy LK, Smith ME, Gevins A (2000) Test-retest reliability of cognitive EEG. Clin Neurophysiol 111: 457-463

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264: 746-748

Mesulam M (1998) From sensation to cognition. Brain 121: 1013-1052

Möttönen R, Krause CM, Tiipana K, Sams M (2002) Processing of changes in visual speech in the human auditory cortex. Brain Res Cogn Brain Res 13: 417-425

Mukamel R, Harel M, Hendler T, Malach R (2004) Enhanced temporal non-linearities in human object-related occipito-temporal cortex. Cerebr Cortex 14: 575–585

Muller UA (1993) Statistics of variables observed over overlapping intervals. November 30, p 10. Working Paper from Olsen and Associates No 1993-06-18. File URL: http://www.olsen.ch/research/403_intervalOverlap.pdf

Näätänen R, Winkler I (1999) The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull 125: 826-859

Neuper C, Pfurtscheller G (1999) Motor imagery and ERD. Event-related desynchronization and related oscillatory phenomena of the brain. In: Pfurtscheller G, Lopes da Silva FH (eds). Handbook of electroencephalography and clinical neurophysiology, Vol. 6, revised ed. Elsevier, Amsterdam, pp 303-326

Nuwer MR, Comi C, Emerson R, Fuglsang-Frederiksen A, Guerit JM, Hinrichs H, Ikeda A, Luccas FJ, Rappelsberger P (1998) IFCN standards for digital recording of clinical EEG. Electroencephalogr Clin Neurophysiol 106: 259-261

O’Hare JJ (1991) Perceptual integration. J Wash Acad Sci 81: 44-59

Perneger TV (1998) What is wrong with Bonferroni adjustments. Br Med J 136: 1236-1238

Pfurtscheller G, Klimesch W (1990) Topographical display and interpretation of event-related desynchronization during visual-verbal task. Brain Topogr 3: 85-93

Pfurtscheller G, Stancak JrA, Neuper C (1996) Post-movement beta synchronization: a correlate of an idling motor area? Electroencephalogr Clin Neurophysiol 98: 281-293

Pöppel E (1994) Temporal mechanisms in perception. Int Rev Neurobiol 37: 185-202

Pöppel E (2004) Lost in time: historical foundations and the 3-second-window of temporal integration. Acta Neurobiol Exp 64: 295-301

Pulvermuller F, Keil A, Thomas E (1999) High-frequency brain activity: perception or active memory? Trends Cogn Sci 3: 250-252

Raij T (1999) Patterns of brain activity during visual imagery of letters. J Cogn Neurosci 11: 282-299

Riley WJ (2003) Techniques for frequency stability analysis. IEEE International Frequency Control Symposium. Tampa, FL, May 4.

Rosenblum LD, Saldana HM (1996) An audiovisual test of kinematic primitives for visual speech perception. J Exp Psychol Hum Percept Perform 22: 318-331

Rothman KJ (1990) No adjustments are needed for multiple comparisons. Epidemiology 1: 43-46

Sakowitz O, Schurmann M, Basar E (2000) Oscillatory frontal theta responses are increased upon bisensory stimulation. Clin Neurophysiol 111: 884-893

Sakowitz O, Quiroga R, Schurmann M, Basar E (2001) Bisensory stimulation increases gamma-responses over multiple cortical regions. Brain Res Cogn Brain Res 11: 267-279

Salmelin R, Hämälinen M, Kajola M, Hari R (1995) Functional segregation of movement-related rhythmic activity in the human brain. Neuroimage 2: 237-243

Sams M, Aulanko R, Hämäläinen M, Hari R, Lounasmaa OV, Lu ST, Simola J (1991) Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. Neurosci Lett 127: 141-145

Schacter DL (1977) EEG theta waves and psychological phenomena: a review and analysis. Biol Psychol 5: 47-82

Singer W (1994) Putative functions of temporal correlations in neocortical processing. In: Koch C, Davis JL (eds). Large-Scale Neuronal Theories of the Brain, MIT Press, Cambridge, pp 201-237

Soto-Faraco S, Navarra J, Alsius A (2004) Assessing automaticity in audiovisual speech integration: evidence from the speeded classification task. Cognition 92: B13–B23

Stampfer HG, Basar E (1985) Does frequency analysis lead to better understanding of human event related potentials? Int J Neurosci 26: 181-196

Steriade M, Gloor P, Llinas RR, Lopes da Silva FH, Mesulam MM (1990) Basic mechanisms of cerebral rhythmic activities. Electroencephalogr Clin Neurophysiol 76: 481-508

Taylor JL, McCloskey DI (1990) Triggering of preprogrammed movements as reactions to masked stimuli. J Neurophysiol 63: 439-446

Thatcher RW (2001) Normative EEG databases and EEG biofeedback. Journal of Neurotherapy 2: 8-39

Tiippana K, Andersen TS, Sams M (2004) Visual attention modulates audiovisual speech perception. Eur J Cogn Psychol 16: 457-472

Treisman AM, Sato S (1990) Conjunction search revisited. J Exp Psychol Hum Percept Perform 16: 459-478

Tononi G, Edelman GM, Sporns O (1998) Complexity and coherency: Integrating information in the brain. Trends Cogn Sci 2: 474-484

Zarahn E, Aguirre GK, D’Esposito M (1999) Temporal isolation of the neural correlates of spatial mnemonic processing with fMRI. Cogn Brain Res 7: 255-268

Yordanova J, Kolev V, Rosso OA, Schurmann M, Sakowitz OW, Özgören M, Basar E (2002) Wavelet entropy analysis of event-related potentials indicates modality-independent theta dominance. J Neurosci Methods 117: 99-109

Usher M, Niebur E (1997) Active neural representations: Neurophysiological data and conceptual implications. In: Peschl M, Riegler A (eds). New Trends in Cognitive Science. University of Vienna, Austria.

Varela F (1995) Resonant cell assemblies: A new approach to cognitive functions and neuronal synchrony. Biol Res 28: 81-95