PART VII

Specialized Methods

28

Neuroimaging Analysis I: Electroencephalography

Josep Marco-Pallarés, Estela Camara, Thomas F. Münte and Antoni Rodríguez-Fornells

COGNITIVE NEUROSCIENCE AND NEUROIMAGING TECHNIQUES

Cognitive neuroscience has been termed the *biology of mind*. As such it is, to a large extent at least, a science about the *human* mind, as many of the higher cognitive functions, including language processing, episodic memory and executive functions, can best or exclusively be studied in human subjects. To fulfil its promise, cognitive neuroscience is in need of techniques that can serve as windows to the brain as it carries out the processes that make up the mind. Since human participants are under study, these techniques need to be non-invasive.

In light of this, the recent success of cognitive neuroscience can be attributed to two factors: the increasingly sophisticated experimental designs that are borrowed from cognitive science and psychology, and, the methodological developments in neuroimaging techniques.

In the present Chapter and the following one we will concentrate on the two major and most widely used neuroimaging techniques, namely methods derived from electroencephalography (EEG) and functional magnetic resonance imaging (fMRI). While EEG has been around for about 80 years, recent methodological advances in signal analysis have led to a renewed interest in EEG-based experiments. Functional MRI, while having a much shorter history of little more than 15 years, has already reached a high level of sophistication, but more developments regarding analysis techniques are to be expected.

For space reasons, we will not discuss other neuroimaging techniques here, but would like to point out that each of these possess unique properties that make them valuable tools in cognitive neuroscience. Near infra-red spectroscopy (NIRS) uses near-infra-red light to non-invasively measure changes in the concentration of oxygenated (O_2Hb) and deoxygenated (HHb) hemoglobin. Light from the near-infra-red spectrum can penetrate the skull and reaches the underlying cortex, where it is partly absorbed and partly reflected.

[15:53 15/4/2009 5283-Millsap-Ch28.tex] Job No: 5283 Millsap: The SAGE Handbook of Quantitative Methods in Psychology Page: 677 675–696

From the amount of reflected near infra-red light, it is possible to calculate changes in the concentration of O_2Hb and HHb. The main advantage of NIRS is that it can be used in participants that are not able to perform tasks in a MRI scanner (e.g., infants, severely compromised patients), and moreover with tasks that could not be performed in a scanner (such as pointing, object manipulation, and so forth). Multichannel systems can be used to provide data with reasonable spatial resolution. We refer the reader to Obrig and Villringer (2003) for a technical description and Horovitz and Gore (2004) for an application to a cognitive neuroscience question.

Positron emission tomography (PET) yields tomographic pictures of the brain based on the decay of injected radioactive tracers. Whereas PET studies of task-related changes in blood flow (using ^{15}O-labeled water or butanol, for example) have mostly been replaced by fMRI, PET gains increasing importance in cognitive neuroscience because of its ability to map neurotransmitter changes during cognitive and other tasks (see Monchi et al., 2006 for an application), and the density of Alzheimer disease plaques (see Cohen, 2007, for a review of the technique).

Whereas transcranial magnetic stimulation (TMS) might not be considered a neuroimaging technique in the strict sense, the ability to create virtual lesions in normal human participants has great potential, in particular when combined with other neuroimaging techniques such as event-related brain potentials (ERPs) (see Rollnik et al., 2004, for an example) or fMRI (see, Ruff et al., 2008, for an example).

SPATIAL AND TEMPORAL PROPERTIES OF ELECTROENCEPHALOGRAPHY AND (FUNCTIONAL) MAGNETIC RESONANCE IMAGING SIGNALS

When groups of neurons are involved in information processing, they show a change in their firing rate. The physiological phenomena associated with this change can be detected and recorded by several neuroimaging techniques, i.e., EEG, magnetoencephalography (MEG), fMRI or PET. Of these, EEG and fMRI are currently the most widely used brain imaging techniques. EEG is usually recorded using between 16 to 128 electrodes that are placed on the intact scalp, whereas fMRI provides information about the hemodynamic response of several thousands of voxels into which the brain is divided. The spatial information of fMRI (in the order of few *mm*) is therefore much better than the one obtainable by EEG (*cm*). This disadvantage of EEG is balanced by its superb temporal resolution (milliseconds), which compares to several seconds in fMRI.

Whereas the information provided by fMRI and EEG respectively may in some sense be viewed as complementary, with fMRI answering the 'where' and EEG the 'when' question in neural processing, it must be cautioned that there is no directly established relationship between fMRI and EEG signals (Logothetis et al., 2001). While both signals are very different in nature and in their temporal and spatial properties, they lend themselves to treatment by similar mathematical and statistical methods, because: (1) the spectral (1/f) behavior of these signals indicate the participation of neural activity on different scales; (2) they both require extraction of the task-related signal from background activity and noise (i.e., noise of the recording device, muscle-activity, heartbeat, head or eye movements); and (3) the experimental designs used in cognitive neuroscience are similar in EEG and fMRI. The latter point is particularly true since the introduction of event-related designs in fMRI studies. In the following sections, we will illustrate the different analytical approaches used in EEG signal [see also Chapter 29 (in this volume, Camara et al.) for MRI analysis and those methods of analysis which are common to both techniques, e.g., independent component analysis].

ELECTROENCEPHALOGRAPHY AND EVENT-RELATED BRAIN POTENTIALS

Basic designs

The basic experimental approach in using EEG and ERPs in cognitive neuroscience is, in principle, not different from that in other areas of cognitive science. Rigid control of participant behavior is usually required, and care must be taken to isolate the cognitive process under study by the experimental manipulation. There are a few aspects, however, in which basic designs of EEG/ERP experiments differ from experiments elsewhere in cognitive science or experimental psychology. First, owing to the low signal-to-noise ratio of single trial ERPs, responses from multiple single trials need to be averaged together. Depending on the size of the component under study, a minimum of 10 (e.g., in the case of the error-related negativity) to up to several hundred (e.g., in the case of selective attention effects) single trials need to be averaged together. To generate enough trials to yield a reliable and robust ERP might not be problematic in most cases, but it can be a limiting factor in other areas. For some experiments in psycholinguistics there are simply not enough stimuli available (see Weyerts et al., 1997, for an example).

A major advantage of the ERP approach is that it is possible to study responses to stimuli to which no overt behavioral answer is required. We will illustrate this by two examples taken from the areas of selective attention and language processing. Consider a typical selective attention ERP experiment like the following: the participant is required to look at the center of a video monitor. Left and right of the fixation point, random series of blue and red bars appear at a rate of about three stimuli per second, most of a certain height with a few just slightly taller. The participant's task is to attend to a particular class of stimuli (e.g., the red bars on the left) and to respond to the rarely occurring slightly taller bars by button press. In this situation, we are able to investigate the attentional filter processes for the stimuli: indeed, ERPs to all stimuli on the attended side of the display show signs of (spatial) attentional enhancement. Those stimuli that share both location and color, but not height, with the target stimulus are associated with an additional selection negativity signifying selection of the color feature (see Hillyard and Münte, 1984, for a full description of the experiment). Importantly, this information about the hierarchical selection implemented in the human brain would not be available with purely behavioral measures.

In the domain of language research, participants are often required to read materials in order to perform a certain (mock) task. Such a task might entail that participants need to answer certain questions on the materials during the break between experimental blocks. Unbeknownst to the subjects, the materials are manipulated in a certain way. Consider for example the following materials (taken from Matzke et al., 2002):

(1) Die begabte Sängerin entdeckte den talentierten Gitarristen.
The gifted singer$_{(Fem.Nom.?Acc.?)}$ discovered the talented guitar player$_{(Masc.Acc.)}$.

(2) Die begabte Sängerin entdeckte der talentierte Gitarrist.
'The gifted singer$_{(Fem.Nom.?Acc.?)}$ discovered the talented guitar player$_{(Masc.Nom.)}$.'
Meaning: The talented guitar player discovered the gifted singer.

In both sentences, the first noun phrase (die begabte Sängerin) is identical but case ambiguous. It could be nominative (as in (1)) and thus serve as the subject of the sentence, or accusative case (as in (2)) and thus serve as the object. Importantly, in German, the dis-ambiguation of the sentence takes place only at the second noun phrase (der/den talentierte/n Gitarristen) but both versions of the sentence are perfectly grammatical. By studying the ERPs to these sentences in a word-by-word fashion, it is possible to glean information about syntactic processing in the brain without directing participants' attention to the different grammatical constructions.

Table 28.1 Types of stimuli in the attention experiment

	Location	*Color*	*Height*
Left/red/tall	+	+	+
Left/red/small	+	+	−
Left/blue/tall	+	−	+
Left/blue/small	+	−	−
Right/red/tall	−	+	+
Right/red/small	−	+	−
Right/blue/tall	−	−	+
Right/blue/small	−	−	−

Note that functional imaging with fMRI shares some of these advantages. Indeed, an fMRI study using the same materials has been performed (Bahlmann et al., 2007).

The experimental examples discussed so far have (implicitly) made use of the subtraction logic first introduced to psychology by the Dutch scientist Donders. Indeed, such logic underlies many ERP and fMRI studies. Consider the attention experiment mentioned above. Again, we examine the situation in which the tall red bars on the left are attended. The other types of stimuli can be classified as shown in Table 28.1.

By rotating attention conditions, several different ERPs can be recorded for each particular stimulus type. The effects of attentional selection by location, color and size can thus be obtained by subtraction of the different ERPs. It has, however, been pointed out that such logic assumes that there is no interaction between the different processes under study, an assumption that is not true in every case. Alternatively therefore, factorial designs may be employed (see Osman, 1998; Sternberg, 1998).

Standard statistical analysis of electroencephalography event-related brain potentials (time-domain approach)

Event-related potentials can be thought of as minute voltage fluctuations that are buried within an ongoing EEG. Therefore, ERPs benefit greatly from signal averaging to enhance their signal-to-noise ratio (SNR). To this end, biosignals are digitized at a fixed rate (for cognitive ERPs 150 to 1000 points per second and channel are usually recorded). Together with the EEG trigger events (related to the onset of a stimulus, a response, or a movement), are recorded. The events of interest are repeated and a time-locked signal average is then calculated across the trial epochs for each time point of the epoch. Formally, this can be expressed as follows: if $X_j(t)$ represents the voltage at a particular electrode at time t and trial j, the signal average is defined as:

$$X_t = \frac{1}{J}\sum_{j=1}^{J} X_{jt} \tag{1}$$

Usually, X_{jt} is considered the sum of signal of interest S_t plus random noise N_{jt} (background EEG and measurement error). Using this method, signal averaging improves the SNR. However, note that this view, as discussed below, might not be entirely true as it seems that at least some parts of the ERP are brought about by phase resetting of the ongoing EEG. The signal power $\hat{\sigma}_s^2$, noise power $\hat{\sigma}_N^2$, and SNR can be estimated using:

$$\hat{\sigma}_s^2 = \frac{1}{T}\sum_{t=1}^{T}\bar{X}_t^2 - \frac{1}{J}\hat{\sigma}_N^2$$

$$\hat{\sigma}_N^2 = \frac{1}{T(J-1)}\sum_{j=1}^{J}\left(\sum_{t=1}^{T}\left(X_{jt} - \bar{X}_t^2\right)^2\right)$$

$$\approx \text{Variable} \quad \bar{X}_t$$

$$\mathrm{SNR} = \hat{\sigma}_S^2 \Big/ \hat{\sigma}_N^2 \tag{2}$$

One of the key assumptions of signal averaging is that the signal is invariant across trials. This is clearly not the case, as it has been shown that some ERP components, such as the P300, vary in a trial by trial manner. If, for example, latency jitter is present for a specific component this will lead to a smearing out of the component in the signal average, and the peak amplitude of the average will thus not properly reflect the component's amplitude in single trials. For certain purposes, realigning the single trials by moving a template (usually the conventional

average or part of a sine-wave) across the single trial epoch and searching for the time-point at which the template and the single trial have the greatest cross-correlation has therefore been tried. This time-point is then used to realign the single trials (see, for example, Wastell, 1977).

Prior to quantifying waveform changes in the average potential, it can be advisable to apply filters that enhance the signal-to-noise ratio for the effects of interest. For example, the error-related negativity (ERN) (see Gehring et al., 1995) has a frequency around 5–6 Hz. To remove contamination by overlapping slow positive waves and by high frequency activity, it might be useful to apply a band-pass filter to remove activity below 2 Hz and above 8 Hz to best bring out the ERN activity.

After the average ERP is obtained for several experimental conditions, say for stimuli in the right visual field while they are attended and for the same types of stimuli when they are outside of the focus of attention, the next analysis step is waveform quantification (for an introduction to standard ERP measures, see Luck (2005) and Picton et al., 2000). The waveforms are characterized by peaks and troughs that lend themselves to quantification (see Figure 28.1 for an illustration). The usual practice is to determine amplitudes relative to a baseline period (e.g., −100 to 0 ms relative to the onset of the event). The voltage of the baseline period is set to 0. Typical parameters that are determined from the waveforms are:

(i) *Peak amplitude*: the most negative or most positive point relative to the baseline is determined within a defined time window.
(ii) *Peak latency*: the latency of the most negative of most positive point within a time window is determined relative to the onset of the time-locking event.
(iii) *Mean amplitude*: the mean amplitude within a given time window is determined; this measure is equivalent to an area measure.
(iv) *Peak-to-peak amplitude*: in cases were several peaks and troughs occur in quick succession it might be adequate to determine the amplitude difference between two successive peaks.
(v) *Onset latency*: the onset latency of a component is notoriously difficult to determine.

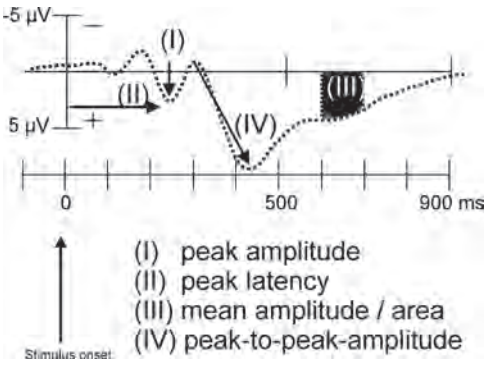

Figure 28.1 Typical parameters determined from the event-related potential waveforms. I. Peak amplitude. II. Peak latency. III. Mean amplitude/area. IV. Peak-to-peak amplitude. Refer to color plates at the end of this volume for a colored version of this figure.

Several suggestions have been made to give an estimate of the onset latency. For example, it might be estimated by determining the time-point at which the amplitude of the rising flank of the component has reached 15% (or some other fraction) of the peak amplitude. This is known as *fractional amplitude latency.* Alternatively, *fractional area latency* might be determined.

Sometimes, the determination of the onset latency is problematic because of residual noise in the waveform. Miller et al. (1998) have therefore suggested a 'jack-knifing' method based on measuring the difference in the onset latencies of two experimental conditions. While this method has been suggested for the measurement of the lateralized readiness potential (LRP), it can readily be applied to other components as well (see Banfield et al., 2006, for an example).

Before measurements are taken of waveforms, it is sometimes useful to perform *waveform subtraction* to reduce the effects of component overlap and to bring out the effect of a specific experimental manipulation. Consider, for example, a selective attention

paradigm in which stimuli in the right visual field are attended in one condition and not attended in another. If the waveforms obtained in the unattended condition are subtracted from those in the attended condition, the resulting *difference wave* presumably reflects the attention effect proper, and quantification of the difference wave should therefore provide a direct measure of the neural correlates of selective attention. However, the computation of difference waves is not a good idea in every case. Consider the situation in which a peak is shifted in latency from one condition to the next, where subtraction will introduce a 'ghost-component' into the difference wave that will be misleading in the interpretation.

One important advantage of the ERP technique is that the resulting data are multidimensional, and that the spatial distribution of effects can be taken into account. Indeed, measurements are usually obtained from several electrode sites, and the resulting datasets lend themselves to statistical analysis by repeated measures analysis of variance (with electrode-sites being treated as one or more factors). One of the potential strengths of such an approach is that conditions can not only be distinguished by effects at certain scalp sites, but also by a differential distribution of an effect across multiple scalp sites. The latter would be reflected by a condition *x* electrode-site interaction in an analysis of variance. In such a situation, one might be inclined to assume that neural generators that are at least partly different might be at work in the two conditions. McCarthy and Wood (1985) have pointed out, however, that there is a fundamental incompatibility between the additive model upon which analysis of variance (ANOVA) are based and the multiplicative effect on ERP voltages produced by differences in source strength. Using simulations, they showed that highly significant interactions involving electrode location can be obtained between scalp distributions with identical shapes generated by the same source. They suggested a scaling method to eliminate overall amplitude differences between experimental conditions before an ANOVA is performed. In other words, McCarthy and Wood (1985) suggested that condition *x* electrode-site interactions that survive vector scaling are indicative of true differences in neural generators between conditions. More recently, however, it has been pointed out by Urbach and Kutas, (2002) that even for ideal distributions of generators and surface potentials, the extent to which vector scaling refines conclusions about generator distributions is limited: 'prior to amplitude normalization, differences in scalp distributions show that neural generators differ in some combination of location, polarity, and relative or overall strength. After amplitude normalization, residual differences merely attest to the fact that neural generators differ in some combination of location, polarity, or relative strength, that is, that they differ in spatial configuration. Of all the possible combinations of differences in generator locations, polarities, and strengths that could account for the different scalp distributions, amplitude normalization at best only rules out one special case: namely, where the generators in the two conditions all have the same locations and polarities and differ in strength by the same multiplicative factor (Urbach and Kutas, 2006). In this sense, these researchers suggest that non-normalized data should be used in assessing condition *x* electrode-site interactions (see further discussion in Luck, 2005).

The measurement techniques discussed so far reveal data about peaks and troughs of the waveform, but not necessarily about ERP *components*. In neurophysiological/psychological terms, a component can be thought of as being generated by a neural or cognitive process, while in a statistical sense a component explains experimental variance (see discussion for conceptual issues regarding ERPs, Rugg and Coles, 1996). Peaks and troughs may thus come about by the superimposition of several components. There have been a number of suggestions for decomposing ERP waveforms in order to isolate their components, among them principal component analysis (PCA) or, more recently, independent component analysis

(discussed in Chapter 29). PCA uses the time points on waveforms from different subjects, different electrodes, and different experimental conditions to define components (for details, see Picton et al., 2000). In statistical terms, PCA identifies orthogonal axes of maximal variance in a multidimensional space defined by the variables. Generally, these axes are rotated according to the varimax procedure, which introduces a certain degree of arbitrariness. PCA solutions are not unique, as many rotations of the factors are possible. Also, the selection of experimental conditions, electrode sites, number of subjects and so on will determine the factor structure of a given experiment. Thus, it is difficult to compare factor structures across experiments and to identify components across different studies.

Statistical analysis of multichannel EEG data is problematic, since the many channels and multiple conditions in one experiment often call for multiple statistical tests, thereby increasing the chance of type I errors (rejecting the null hypothesis when it is true). The best way to circumvent this problem is by replication of initial findings in a second independent group of subjects, i.e., by running a confirmatory study. The typical corrections used to compensate for increased type I error (e.g., Bonferroni type corrections) may over-correct, since data from adjacent electrodes are correlated and not independent. Whereas rigid and theoretically well-grounded methods for statistical corrections have been described in functional imaging and have become standard procedures in that field (see below), such procedures have been less widely performed in ERP research and are far from being standardized.

Artifact rejection and correction algorithms

Prior to further processing and averaging, the EEG has to be checked for undesirable electrical noise and artifacts resulting from movements, eye movements and blinking, and muscular activity. Artifacts may either be rejected (i.e., those stretches of the signal contaminated are removed from further processing) or corrected.

For artifact rejection, the most widely used criterion is to establish a threshold value for artifact amplitude (usually between ± 50 and ± 100 μV). Another common procedure is to reject those trials that present a specified abnormally steep slope or drift. Finally, trials presenting technical problems (such as amplifier saturation) are also removed from further analysis. If participants are not given specific instructions about artifacts or visual feedback about the effects of blinking and moving in the EEG, between 15% and 30% of the trials require rejection. However, if participants collaborate and brief pauses or blinking periods are introduced (if possible incorporated into the design of the task), the rejection rate is about 10%.

Such 'rejection' techniques will be problematic in situations in which the number of trials per condition is very low (e.g., below 25), which is the case, for example, in some psycho-linguistic studies, or in studies with patients or special populations (newborns, children, etc.). Thus, it is necessary in such cases to remove the noise from the contaminated signal in order to be able to use most of the trials in the averaging process. Most commonly used algorithms for cleaning blinks and ocular movement are based on regression analysis. In this type of analysis, the contamination of each electrode by a certain type of artifact is assessed by computing a propagation factor, which is used to remove the estimated signal influence (Verleger et al., 1982). Another approach is to use dipole modeling to isolate the ocular activity (Berg and Scherg, 1991). Although the latter method seems to work reasonably well (see Lins et al., 1993), the most important problem of such approaches is that ocular electrodes might also pick up EEG activity proper and that these methods might therefore also remove part of the signal of interest. Alternatively, it has been proposed that the application of Blind Source Separation techniques to EEG/ERP data can reliable remove ocular, muscular and electrical noise artifacts from the raw signal (Jung et al., 2000).

Source analysis

While temporal information can be readily inferred from a scalp-recorded EEG, the question as to where a particular signal is coming from has been of interest for EEG researchers from the very beginning, first to localize focal (epileptic) activity and nowadays to pinpoint neural structures responsible for particular cognitive operations.

A first important issue in the localization of the sources of the EEG is the use of an adequate number of electrodes. While the recording, storage and analysis of up to 256 electrodes is no longer technically a problem, increasing the number of electrodes will prolong the recording session. Thus, a compromise between the number of electrodes and the time needed to run the experiment is necessary. Lantz et al. (2003) have shown that going from 32 to 64 electrodes markedly changes localization results, but that a further increase of electrode density yields little additional precision.

Electrogenesis of scalp potentials

The electrical potential recorded in the scalp is a consequence of the electrical activity of large assemblies of neurons that are activated synchronously. Although the exact electrogenesis is not fully understood, it is supposed that the activity registered using EEG is related to the influx of positive ions across the post-synaptic membrane when a neurotransmitter is released. In addition, there is a re-distribution of charges in the outer part of the membrane. If several (thousand) neurons depolarize synchronously, the total net current can be recorded at the scalp by using macro-electrodes (see Figure 28.2A and B and Nunez and Srinivasan, 2005).

The surface activity is very dependent on the position and geometry of the neurons involved, as can be seen in Figure 28.2C. Hence, 'closed field' configurations, where neurons are not aligned in parallel may not produce detectable fields at the scalp. In addition, the electric potential presents a very fast decay with distance (see Figure 28.2D), and is further attenuated by the tissues between the source of the potential and the scalp electrode (i.e., skull and scalp). Hence, EEG is dominated by activity from those areas presenting an 'ordered' geometry and relatively close to the scalp. Pyramidal neurons in the neocortex are thus the main generators of EEG signals. They comprise three-quarters of all cortex neurons and can fire synchronously, because of the local density of excitatory interneurons. Whereas electrical activity from subcortical structures is less well detected given the non-ordered nature of the cell assemblies and the long distance to the scalp, activity of some structures such as auditory nerve or some brain stem structures can be detected at the scalp. Because the activity of interest is very small in these cases (less than 1 μV; see Harkins, McEvoy, and Scott, 1979), many hundreds or even thousands of stimulus repetitions are necessary.

The application of the physical rules of electromagnetism (the Maxwell equations) to the brain electromagnetic currents allows establishing two main problems. The *forward problem* states that, given a source and the electrical characteristics and position of the different layers of the brain, we can unequivocally determine the electrical potential generated by this source. In contrast, the goal of an *inverse problem* is to find a solution that is compatible with certain voltage distribution. Unfortunately, infinite solutions exist compatible with a certain voltage map given that this is an ill-posed problem. Although some physiological constraints (see below for implementations) can be reasonably imposed (i.e., sources of EEG can only be generated in the brain), the inverse problem continues to be a challenge.

Scalp current density

A first approach in increasing the spatial resolution of the EEG is the application of the scalp current source density *scalp current density* (SCD) approach. It is based on the application of a 2D Laplacian operator (∇^2) in two dimensions to the scalp EEG potential data. It can be demonstrated that this

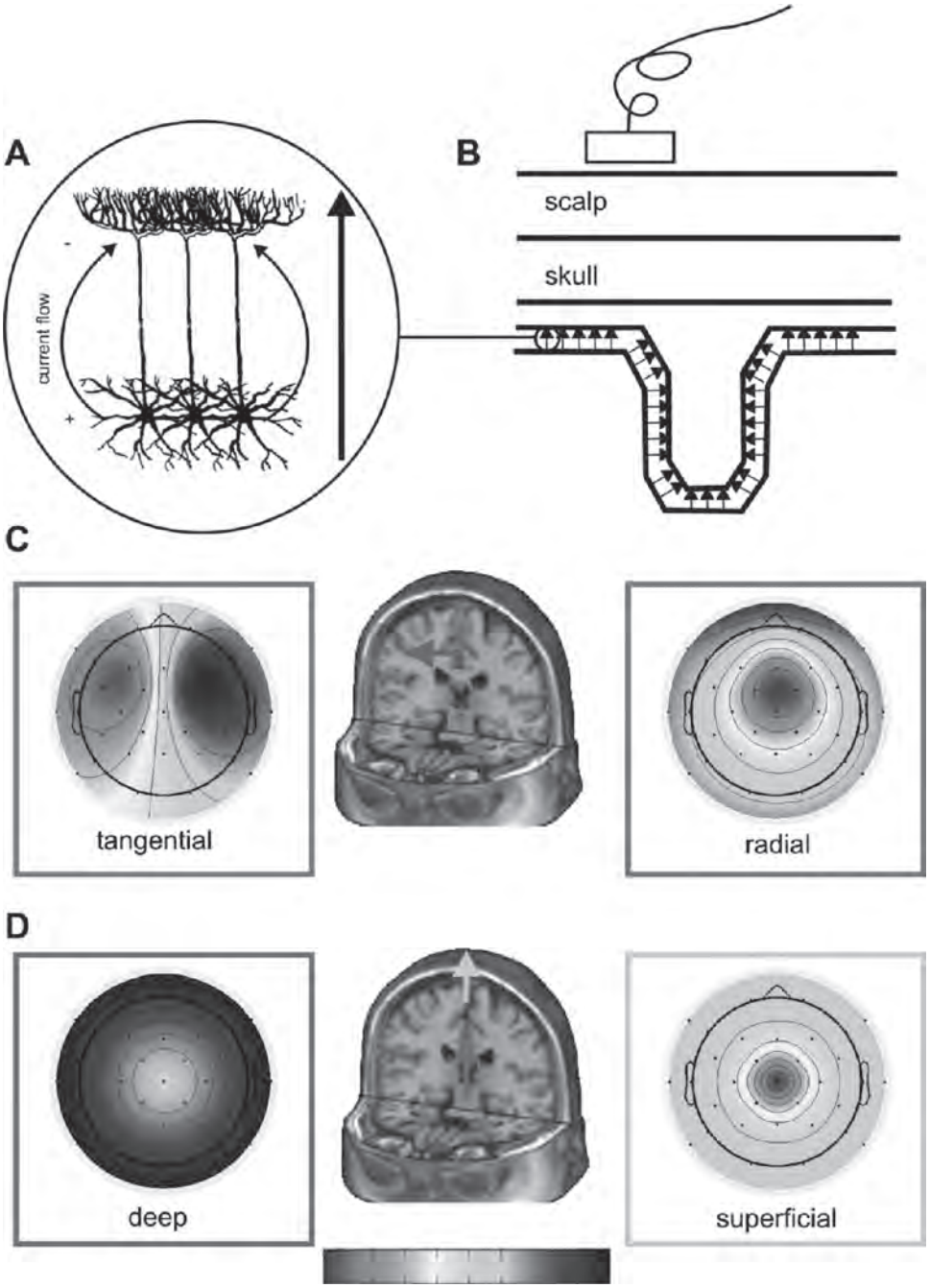

Figure 28.2 Electrogenesis of brain potentials. (A) Current flow generated in a neural assembly of pyramidal neurons. (B) The direction of the current flow in the brain depends on the position of pyramidal neurons in the brain. (C) Voltage activity generated in the scalp by a radial (blue dipole, left scalp voltage) and tangential (red dipole, right scalp voltage) electrical dipole. (D) Voltage activity generated in the scalp by a deep (green dipole, left scalp voltage) and superficial (yellow dipole, right scalp voltage) radial dipole. Refer to color plates at the end of this volume for a colored version of this figure.

measure is equivalent to finding the variation of the normal component of the electrical field, and that it is related to changes in the normal SCD, creating 'sources' and 'sinks' of current density. Another interpretation of the Laplacian operator is that it acts as a spatial high-pass filter, enhancing the spatial resolution but also amplifying the noise of the signal. Several techniques have been proposed for computing the SCD, with the Hjorth method (Hjorth, 1975) and spherical splines (Perrin et al., 1989) being the most widely used.

An example of the advantages of using SCD instead of voltage data can be found in Figure 28.3. Two sources located in the left and right supratemporal cortex generate a midline frontocentral voltage distribution. The application of an SCD algorithm suggests two sources, corresponding to the two internal generators. In spite of the favorable result in this example, the application of SCD can be also very problematic because of noise amplification. Also, it can only provide information about *possible* surface sources, because voltage fields generated by deep sources dissipate and spread across the scalp.

The inverse problem

The inverse problem consists of finding the current density sources that produce a certain voltage. It is inherently ill-posed as there is no unique solution for solving the problem.

Source estimation procedures can be divided in two main groups: dipolar solutions and distributed solutions. In dipolar solutions, the number of sources that are used in the modeling is set *a priori* based on *a priori* constraints (e.g., anatomical or physiological information). Then, this fixed number of dipoles is placed in the brain and their position

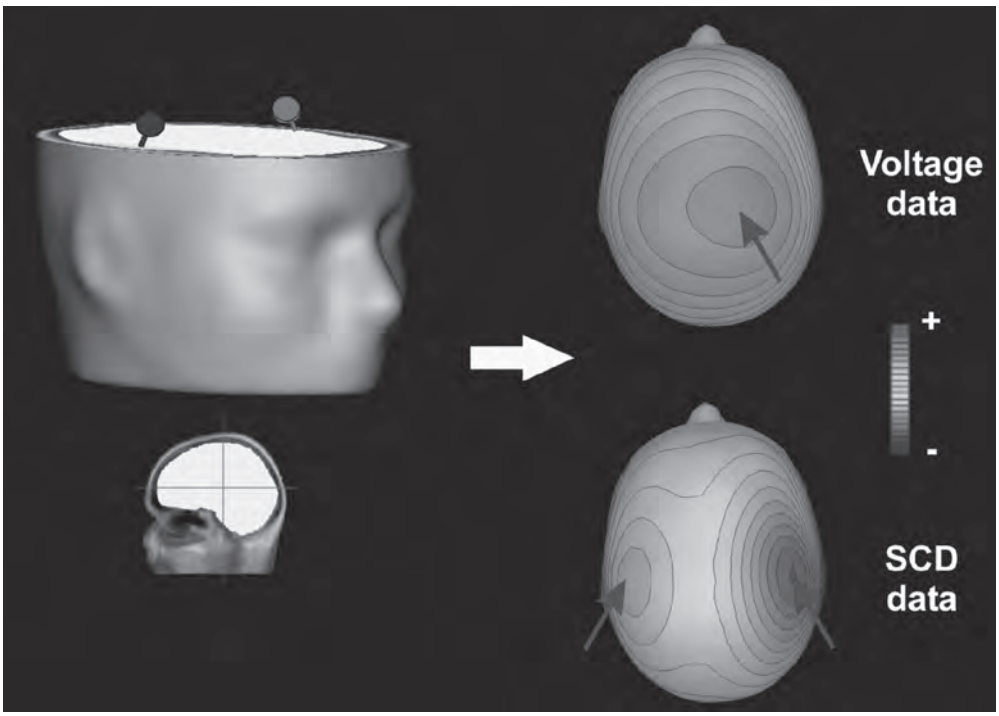

Figure 28.3 Illustration of source current density (SCD) computation. Two dipoles are placed in the scalp at temporal areas. The isovoltage map shows a central positivity. The application of an SCD algorithm distinguishes two sources. As can be seen, the picture given by the SCD is closer to the real source configuration. Refer to color plates at the end of this volume for a colored version of this figure.

and orientation is found by minimizing the difference between the scalp potential produced by the dipoles (computed using the forward solution) and the real scalp potentials. In the distributed models, the head (or only those regions that can reasonably be assumed to generate brain potentials) is divided into a large number of voxels and a solution is found by imposing certain constraints.

Dipolar solutions

Given a certain geometry (location of the different tissues), the voltage produced by a dipole located in any part of the brain can be easily computed (forward model). The first solutions of inverse problems used the forward solution to solve the inverse one. The idea was simple: the number of sources (dipoles) was defined *a priori* and different plausible physiological locations were selected for the dipoles. Also, the orientation of the dipoles could be defined *a priori* based on physiological parameters (i.e., the same orientation as the corresponding pyramidal neurons). The forward problem was computed for all possible solutions, and the voltage map that best explained the real voltage distribution was selected. In addition, the explained variance of the solution gave an estimate of the goodness of fit of the dipolar solution.

Although in use for more than 20 years (Scherg and von Cramon, 1985), dipolar solutions are still popular given their simplicity and the fact that they provide good solutions when few and spatially circumscribed sources are expected that contribute to the observed distribution. They do, however, have a number of limitations. First, the number of dipoles of a solution has to be defined *a priori*. Second, selecting a large number of dipoles increases computation time significantly, making it unfeasible to work with many sources.

Distributed models

Distributed models search for the solution of the inverse problem in a 3D mesh composed by a large number of voxels (generally exceeding the number of electrodes used to register the data). The problem is generally ill-posed, since there are more solutions than equations. There are several different ways of tackling this problem, and the most commonly used are dealt with below.

L2 norm solutions.

These are based on minimization of the modulus of the density vector; in other words, choosing the minimum energy vector. Depending on the choice of the constraints imposed to the problem, solutions can present different characteristics. The most frequently used is the weighted minimum norm solution with low-resolution tomography (LORETA), which is based on the weighted modulus minimization by a Laplacian operator (Pascual-Marqui et al., 1994) resulting in a smoothed solution. This is one of the most widely used methods for EEG source localization.

L1 norm solutions

In L1 norm solutions, the minimization is performed using the L1 norm. One of the most popular applications of the L1 minimum norm is the FOCUSS approach (Gorodnitsky et al., 1995). In general, L1 solutions provide sources less sparse than L2 minimum norm solutions, but their application is difficult because they must be computed recursively.

Other solutions

Other solutions proposed to solve the inverse problem are:

- *Standardized solutions*: based on standardization of the estimation of the currents, given the covariance matrix of estimated noise (Dale et al., 2000; Pascual-Marqui, 2002).
- *Beamformer solutions*: based on spatial filters (Van Veen et al., 1997).
- *Biophysical restrictions based solutions*: i.e., ELECTRA (Grave de Peralta et al., 2000).
- Combination of two different solutions: i.e., shrinking LORETA-FOCUSS (Liu et al., 2004).

The impact of these solutions in the literature is limited at present and therefore they are not discussed in any more detail.

Analysis of the frequency components of electroencephalography

Spectral properties of electroencephalography

A remarkable property of the EEG is its oscillatory behavior. Traditionally, the frequency bands have been divided into delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–25 Hz) and gamma (> 25 Hz) bands. Systematic changes to these EEG bands can be found as a function of behavioral states (i.e., sleep or wakefulness), cognitive tasks, drug intake or neuropsychiatric disorders. The description of such spectral components is therefore a key aspect of EEG studies.

The most widely used spectral analysis approach is based on the fast Fourier transform (FFT). As stated by the Fourier theorem, a signal $s(t)$ can be decomposed as a sum of sinusoidal signals. In the Fourier approach, this can be written as:

$$s(t) = \int_{-\infty}^{\infty} S(\omega) e^{2\pi i \omega t} d\omega \qquad (3)$$

being:

$$S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-2\pi i \omega x} dx \qquad (4)$$

$S(\omega)$ are complex coefficients, whose square gives a measure of the power at the frequency ω. This is the value that has traditionally been used to determine the power of each EEG frequency band.

To illustrate this, in Figure 28.4 we have generated a signal composed from two sinusoidal signals: one with a frequency of 5 Hz, and the other with a frequency of 17 Hz and half the amplitude of the first. In addition, we have added some white noise to the signal (Figure 28.4A and B). The Fourier theorem can then be applied to the resulting signal. The left part of Figure 28.4C shows the result of the complex $S(\omega)$. Two maxima corresponding to the sinusoidal signals are found. To better see these values, we can compute the power spectra of the signal by squaring $S(\omega)$. In this representation, we can clearly see two peaks at the frequencies corresponding to the component signals.

The need for time-frequency approaches

One of the main problems of the FFT and related methods is the fact that temporal information is lost in the computation of the spectral content, which is not always adequate in the study of cognitive functions. Although some states change the global spectral content of the EEG (i.e., the presence of slow waves in the deep sleep is greater than in awake states), their spectral properties may change rapidly in other conditions, i.e., after a stimulus presentation, during the performance of a task, etc. In such situations, short-lived changes in certain frequency bands may occur that need to be detected by adequate methods.

In Figure 28.5 we have illustrated an example of two different trials that present similar responses. At 200 ms there is an alpha band (10 Hz) response and at about 700 ms there is an increase in the beta range (20 Hz) that lasts 100 ms. The FFT analysis (Figure 28.5B) shows a global increase in the 5–20 Hz range but there is neither a clear delineation of the frequencies involved nor any information on the timing. However, when performing a time-frequency analysis, a clear enhancement of the alpha band from 100 to 200 ms and of the beta band from 700 to 800 ms is seen. Hence, the information provided by the time-frequency approach is richer and more appropriate for cognitive neuroscience applications.

Figure 28.5 also shows another important aspect of the time-frequency approach. The two examples are not only different in their spectral content (10 and 20 Hz, respectively), but also in their phase. As can be observed, the activity in the first example is phase locked; that is, its peaks and valleys coincide in time with regard to the stimulus. This results in a clearly visible ERP (mean of the single trials), a type of response known as an *evoked response*. While the increase in power occurs at similar time points, the responses are not phase locked in the second example (the vertical red line coincides with a peak in the first trial and a valley in the second

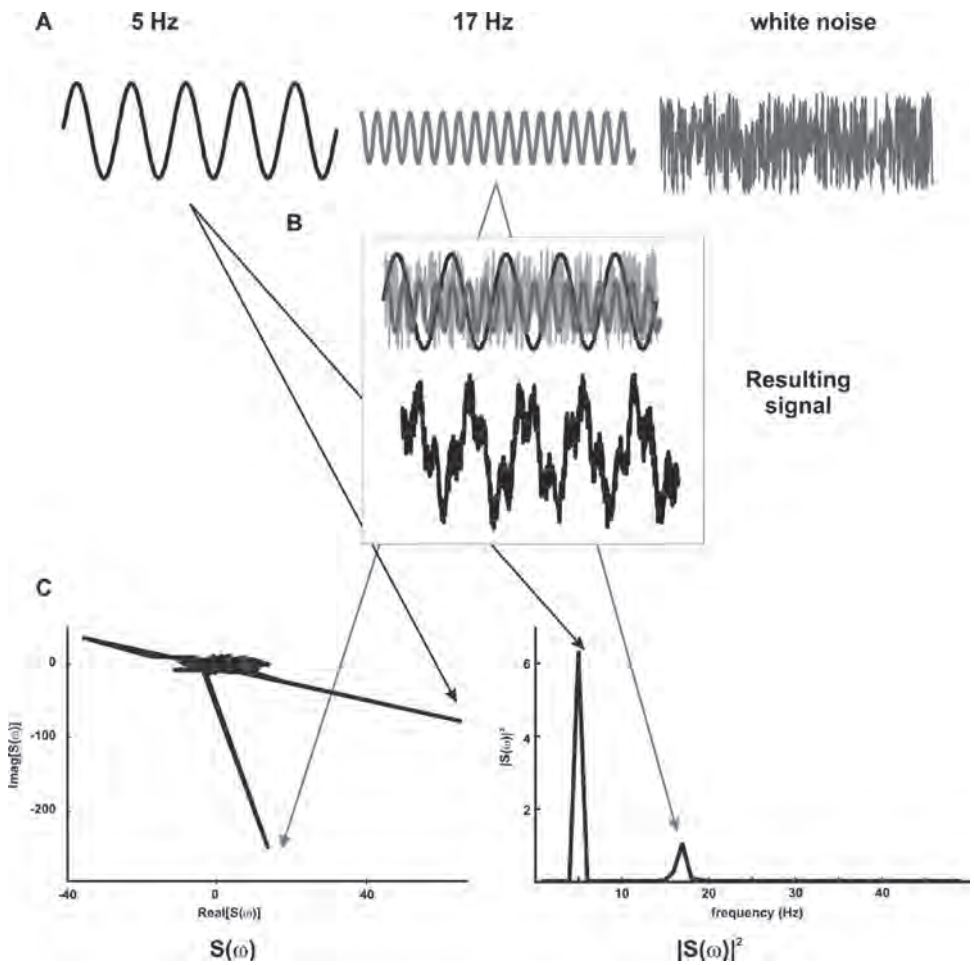

Figure 28.4 Fourier transform computation. (A) A signal is created combining three original signals: a 5-Hz signal, a 17-Hz signal and random noise. (B) By adding the two sine wave signals and the noise, a combined signal is obtained. (C) The left panel shows the complex signal of the fast Fourier transform (FFT) applied to the data. Each complex point in the graph is associated with a specific frequency. Two main frequencies are retrieved with peaks at 5 and 17 Hz. The right panel shows the square of the magnitude of FFT at each frequency. Again, two main peaks reflect the frequencies of the initial signals. Note that square of the magnitude of the 5-Hz signal is greater than the 17-Hz signal, as in the original signals. Refer to color plates at the end of this volume for a colored version of this figure.

trial). Hence, in a time-domain average, the responses are partially or totally cancelled and therefore do not contribute to the ERP (or produce only a very small signal, as in Figure 28.5). Note that this problem is usually greater for fast oscillations than for low frequency oscillations, as well as for oscillations with a longer latency with regard to the stimulus onset.

However, the mean of single trial time-frequency decompositions does show a response that it is not affected by the cancellation effect. Such non-phase-locked responses are known as *induced responses.* The detection of induced responses requires single-trial time-frequency analysis. The different nature of evoked and induced responses also underscores the importance of studying

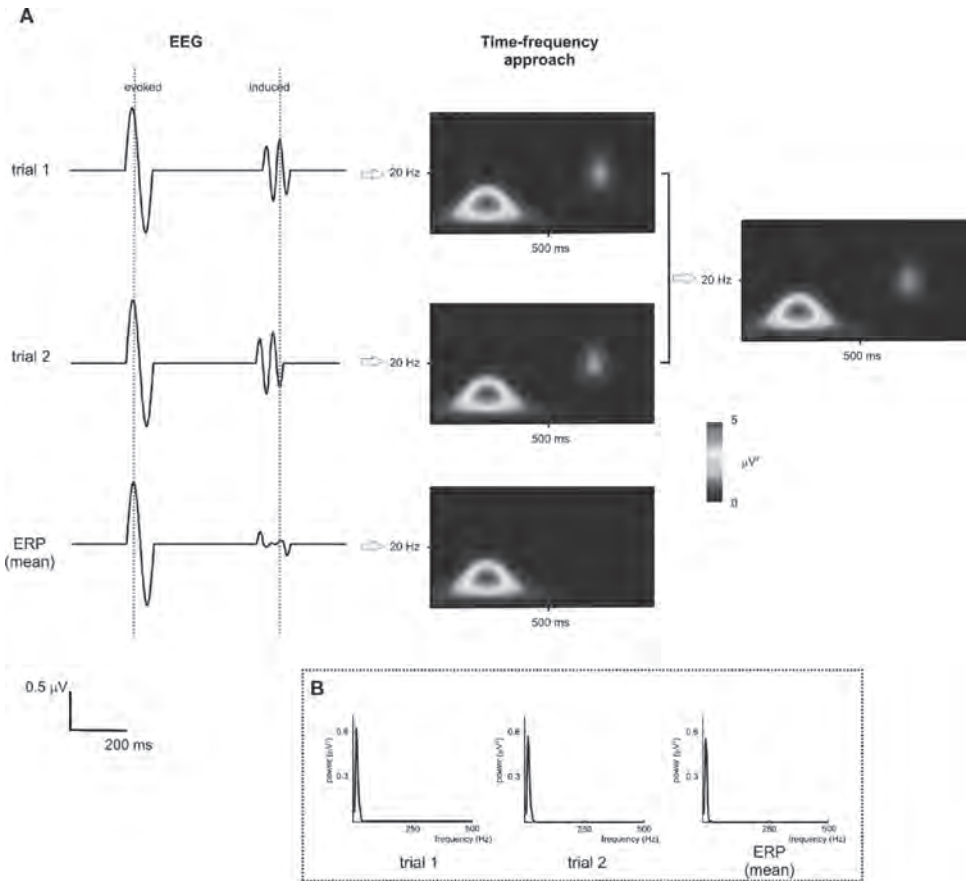

Figure 28.5 Trial by trial time-frequency decomposition. (A) Two trials are depicted with two signals in each: a 10-Hz signal time-locked with respect to the stimulus onset (evoked activity) and a 20-Hz non-phase-locked signal (induced activity). The average of the signals is observed in the corresponding ERP waveform. In this average waveform, the 20-Hz signal has almost been abolished in the time-domain average. The time-frequency decomposition of the average ERP waveform (right) shows only the increase at the 10-Hz but not at 20-Hz signal. In contrast, the time-frequency decomposition of each single trial shows both signals, and this information remains when both decompositions are subsequently averaged together (right panel). Hence, induced activities cannot be studied by applying time-frequency analysis to the averaged ERP responses. (B) The FFT similarly fails to detect both signals. Refer to color plates at the end of this volume for a colored version of this figure.

the phase of the signal. Indeed, some studies have demonstrated that the phase of the EEG can be altered during cerebral processing (Fuentemilla et al., 2006; Makeig et al., 2002)[1]. Given that the phase of the signal can only be studied effectively in single-trial data, a complete time-frequency study should involve time-frequency trial by trial computation of any increase/decrease of power and phase alignment of the signal. To quantify the degree of phase alignment, a measure referred to as inter-trial coherence (ITC) (Makeig et al., 2002) or phase-locking factor (Tallon-Baudry et al., 1996) is used:

$$ITC = \frac{1}{n}\left|\sum_{i=1}^{n}\frac{S_i(\omega)}{|S_i(\omega)|}\right| \qquad (4)$$

where $S_i(\omega)$ is the coefficient at frequency ω of the i-th trial. ITC gives a measure of the degree of similarity of phase over different trials. It ranges from 0 to 1, with 0 indicating

a randomly distributed phase over different trials, and 1 a perfect match of phase in different trials.

Application of frequency analysis to electroencephalography data

In the following section, we will describe three of the most widely used methods: (1) the classical method for computing event-related synchronization/desynchronization (ERS and ERD) based on filtering the data at different frequency bands; (2) the short FFT method, in which the data is divided into windows of identical length to which FFT is applied; and (3) the wavelet analysis, where the same mother wavelet is contracted or dilated to variable length windows.

The first proposal to analyze changes in the power of brain electrical signals was made by Pfurtscheller and Aranibar (1977) to assess increases (ERS) and decreases (ERD) of power in certain frequency bands as compared to baseline (see Figure 28.6C): (1) the EEG single trial data is first filtered at the selected frequency band; (2) the amplitude is then squared and the mean of all trials is computed; finally (3) the percent of increase/decrease of power with respect to baseline is computed.

In the short fast Fourier Transform (SFFT), an FFT is applied to successive short time intervals (see Figure 28.6D). The original signal is convoluted by using sinusoidal signals in a fixed temporal window, obtaining estimations of the power and phase at different frequencies and time ranges. The problem comes from the use of a fixed time window to compute the FFT: if good temporal resolution is needed, a short window is required, whereas a good frequency resolution requires long windows. In other words, there is a trade off between the temporal and frequency resolution, as short time windows comprise less cycles for the sine signals and hence lead to bad spectral resolution. Increasing the length of the windows, on the other hand, decreases temporal resolution. This problem has limited the use of SFFT in EEG analysis.

Finally the most widely used method is wavelet analysis (illustrated in Figure 28.6E), in which the signal is not convoluted by a sine or a cosine, but by a certain signal, namely the mother wavelet. The shape of this signal is the same for all frequencies, but the mother wavelet is expanded or contracted depending on the frequency studied (based on a certain parameter, called the scale a). The convolution is performed at all the time points by moving the wavelet across time using a certain latency shift b, which led to the name *continuous wavelet transform* (CWT). Mathematically, the convolution between the signal $s(t)$and the mother wavelet $\psi^{(}_{b,a}t)$can be written as:

$$W_{\psi_{b,a}}s(t) = \int_{-\infty}^{\infty} s(t)\psi^{*}_{b,a}(t)dt \qquad (5)$$

being $\psi(t)$ the mother wavelet and $\psi_{b,a}(t)$:

$$\psi^{(}_{b,a}t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \qquad (6)$$

Several mother wavelets can be used in the computation of the EEG. The most widely used are the morlet and complex morlet wavelets that comprise a sinusoidal function enveloped by a Gaussian.

As can be seen in Figure 28.6D and E, there are some differences between SFFT and wavelet analysis: in short FFT the length of the windows used is constant, whereas in wavelet analysis they change with frequency. While in short FFT the number of cycles of the sinusoidal signal changes with frequency, the shape of the wavelet is always the same in wavelet analysis. However, in the latter method, there is an inverse relationship between the length of the window in the time and frequency domains: when the frequency increases, the length of the window is shorter in the time domain, but larger in the frequency domain.

One common aspect of all three methods is that after the application of the particular algorithm, data are squared to avoid cancellation when averaging different trials and to convert complex numbers (i.e., in complex morlet wavelets) to real power. For most applications, power changes need to be related to a certain baseline. Thus, the interest is not

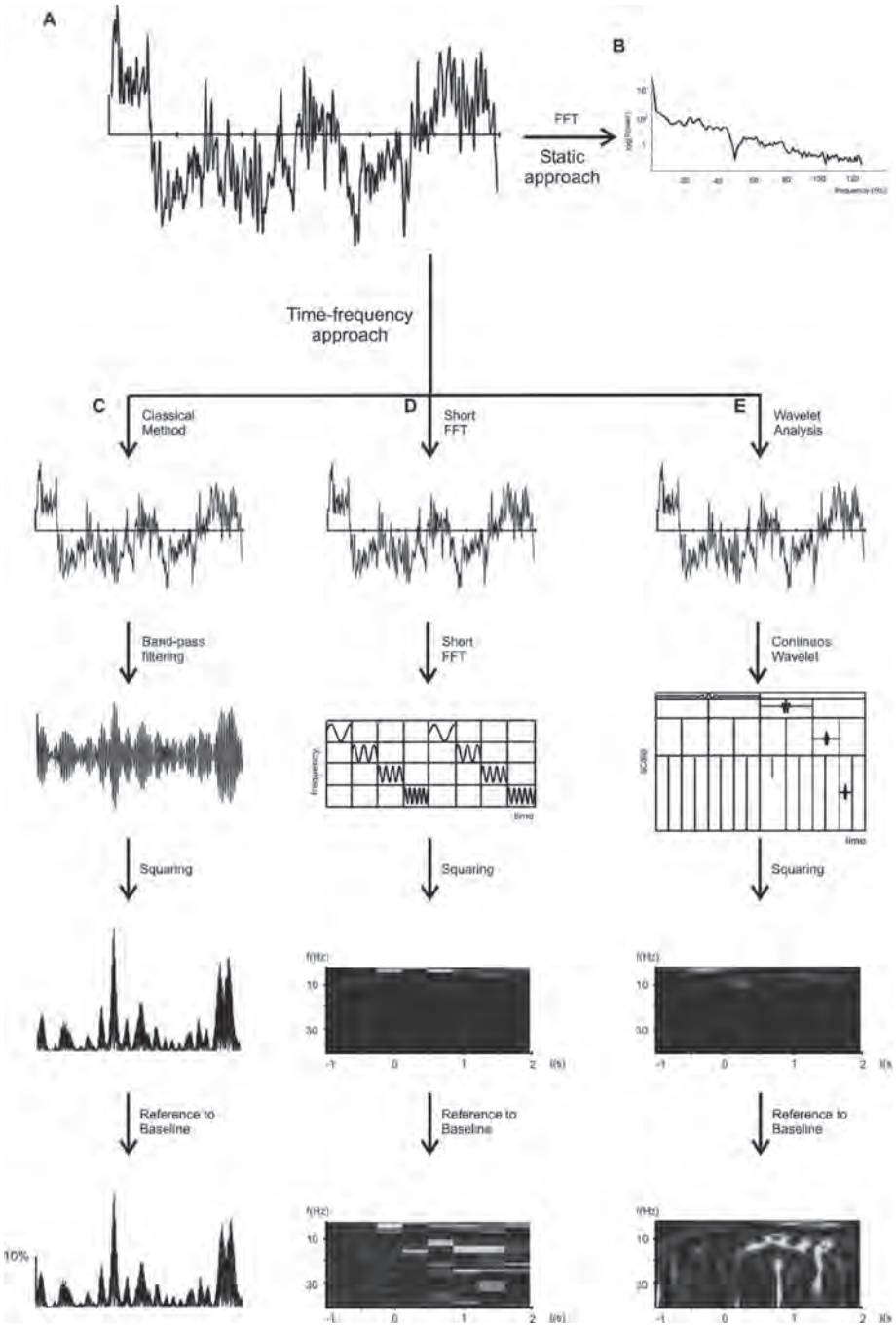

Figure 28.6 Illustration of methods for analyzing the spectral content of a brain electrical signal. (A) A real electroencephalogram (EEG) waveform is used for the analyses. (B) The static approach using a fast Fourier transform (FFT) provides no temporal information. The signal presents a typical 1/f decay, with a decrease at 50 Hz due to a notch filter applied to the data. (C) Classical method for time-frequency decomposition. The data is band-pass filtered, squared and referenced to a baseline. (D) Short-time FFT. An FFT is applied in fixed temporal windows. Then the data is squared and referenced to baseline. (E) Wavelet approach. A wavelet analysis is applied. In this analysis, the temporal and frequency windows vary their length as a function of the frequency. Finally the data is squared and referenced to baseline. Refer to color plates at the end of this volume for a colored version of this figure.

in the absolute value of the power at a certain frequency, but its task- or state-related relative increase or decrease. Referencing the baseline also avoids problems due to the 1/f behavior of EEG data, which generally leads to greater power at lower frequencies (see, FFT in Figure 28.6B).

Statistical analysis of electroencephalography oscillatory activity

While the power computed by spectral techniques follows a χ^2 distribution (Kiebel, et al., 2005) have demonstrated that certain transformations of the spectral data allow the use of the general linear model (GLM) for analysis. One way of performing tests is to compute the average in a spectral and temporal window. The central limit theorem ensures that these averages will follow the Gaussian assumptions. Averaging over several trials is similar. For single trial analysis, log or power transformations generate distributions close to Gaussian. The conclusion of Kiebel et al. (2005) therefore was that EEG power measures can be analyzed using parametric statistics.

A different approach has to be taken with the analysis of phase. Typically, phase is not analyzed individually but by using ITC (Equation 17), that is a measure of the degree of the coherence between different trials. Statistical analysis of such data can be performed by using bootstrap analysis (see Makeig et al., 2002), or non-parametric tests (i.e., U-Mann Wilcoxon or Kruskal–Wallis test). Although ITC follow a Raleigh rather than a normal distribution, some studies have nevertheless used parametric tests (ANOVA).

Finally, it has to be noted that time-frequency analysis suffers from a severe multiple comparisons problem. Imagine a study with 64 electrodes, with epochs of 600 ms length at 500 Hz sampling rate (300 points) and frequencies studied from 1 to 45 Hz. If comparisons are performed point by point, this yields 864,000 possible tests (64 electrodes x 300 time points x 45 frequencies). This is very similar to the number of comparisons performed in fMRI studies. However, in contrast to fMRI, the multiple comparison problem in time-frequency analysis has received only limited attention.

To avoid the multiple comparisons problem, many time frequency studies focus their statistical comparisons on certain spectral-temporal windows and electrodes where the desired effects seem to be present. This *a priori* approach reduces the number of comparisons and possible false positives, and hence, the problem associated with them. A more straightforward solution is based on non-parametric permutation tests (Maris and Oostenveld, 2007). In this approach, the two conditions are compared by means of a standard statistical test (i.e., *t*-test), and points presenting statistical values greater than a certain value are selected. Clusters are created by joining adjacent points on a temporal, spectral and spatial (i.e., electrodes closer than 4 cm) basis. Then, a *cluster-based statistic* is computed on the sum of the values of the statistical test used in a cluster, and the maximum of these values among the different clusters is selected. The next step is the creation of a large number of random partitions by randomly assigning single trials to the 'conditions'. The statistical test is computed for each random partition, and in the points of the cluster presenting the maximum *cluster-based statistics*. This creates a histogram of the random partitions. Finally the proportion of random partitions that result in larger test statistics than the *cluster-based statistics* previously found is computed (*p*-value). Clusters presenting a *p*-value smaller than the critical alpha level (usually 0.05) are then accepted as presenting significant differences between conditions.

CONCLUSIONS

In the last 30 years the utilization of EEG signals for studying brain functions has suffered an important impulse. EEG measurements have evolved from pure clinical settings (i.e., diagnostic of epilepsy, polysomnography, exogenous evoked potentials, etc.)

to become one of the most non-invasive techniques used to the study brain functions and cognitive processes. In addition, it is the only technique (with MEG) that allows non-invasively studying brain functioning at the sub-second temporal domain. One important reason for this reborn has been the incorporation of EEG to the study of cognitive functions, such as language and executive functions. These domains were traditionally studied using behavioral data, but the incorporation of new experimental paradigms has allowed investigating the electrical signatures and oscillatory changes related to these functions. In this regard, several discoveries were critical in the development of the research in these fields, as for example, the N400 component (related to semantic analysis), P600 (associated to syntactic processing) or the error-related negativity (associated to the detection of erroneous responses). At the same time, and more recently, the advent of new techniques of analysis has allowed a richer interpretation of the results and has opened a door for new ways of studying electrical responses associated to cognitive functions. Therefore the analysis of brain oscillations and their properties (changes in power and phase) by means of time-frequency analysis is today very important in order to have an accurate description of brain functioning and brain dynamics. In addition, the possibility of localizing the neural sources of brain electrical activity allows better interpretation of results and the confirmation of existing brain-wired theories or neural-constrained cognitive models. However, the information provided with localizing techniques has to be always interpreted cautiously and it has to be confirmed using complementary techniques that have a better spatial resolution, as for example fMRI (see Chapter 29, in this volume).

In this regard, a promising future for EEG is to combine its information with other functional techniques (such as fMRI, PET or TMS) and the application of new algorithms and techniques in order to extract more information from the raw data. However, the crucial point will be the creation of new research paradigms that allow studying psychological functions in today still emerging fields of cognitive neuroscience. Hence there is a need of new paradigms that allow the application of electroencephalographic techniques to social and developmental psychology, as well as to single-trial experiments. Only the combination of smarter paradigms and powerful techniques of analysis will allow us to face the new challenges of psychology and cognitive neuroscience.

NOTE

High resolution color figures of this chapter can be found at www.brainvitge.org/HQMP

NOTES

1 The study of phase has become increasingly relevant due to the controversy over the origin of evoked potentials. The classical (evoked) theory supports that the ERP arises from a fixed latency fixed polarity response that appears in the EEG (acting as noise) *de novo*. As an alternative, it has been proposed that the ERPs appear due to a reorganization in the phase of the EEG background signal, that is consequently not regarded as noise but as containing relevant information that might affect the EEG response (oscillatory model). Some studies using real data have suggested that both processes might contribute to the ERP (Fuentemilla et al., 2006), making the study of phase important in EEG analysis. However, Yeung et al. (2004) proposed that current methods cannot dis-ambiguate the question about the origin of the ERPs. These authors argued that finding an increase of the inter-trial phase coherence in parallel to the appearance of an ERP (power increase) does not fully support the oscillatory model, because this effect could also be explained by the presence of a fixed latency and polarity response as proposed by the classical model.

REFERENCES

Bahlmann, J., Rodriguez-Fornells, A., Rotte, M. and Münte, T.F. (2007) 'An fMRI study of canonical and noncanonical word order in German', *Human Brain Mapping*, 28: 940–949.

Banfield, J.F., van der Lugt, A.H. and Münte, T.F. (2006) 'Juicy fruit and creepy crawlies: an electrophysiological study of the implicit Go/NoGo association task', *Neuroimage*, 31: 1841–1849.

Berg, P. and Scherg, M. (1991) 'Dipole modelling of eye activity and its application to the removal of eye artifacts from the EEG and MEG', *Clinical Physical and Physiological Measures* 12: Suppl A, 49–54.

Camara, E., Marco-Pallares, J., Münte, T.E. and Rodriguez-Fornells, A. (2009) 'Neuroimaging analysis (II): Magnetic Reasonance Imaging', in Millsap, RE. and Maydeu-Olivares, A. (eds.), *Handbook of Quantitative Methods in Psychology*. SAGE.

Cohen, R.M. (2007) 'The application of positron-emitting molecular imaging tracers in Alzheimer's disease', *Molecular Imaging and Biology*, 9: 204–216.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D. and Halgren, E. (2000) 'Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity', *Neuron,* 26: 55–67.

Fuentemilla, L., Marco-Pallares, J. and Grau, C. (2006) 'Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials', *Neuroimage*, 30: 909–916.

Gehring, W.J., Coles, M.G., Meyer, D.E. and Donchin, E. (1995) 'A brain potential manifestation of error-related processing', *Electroencephalograpy and Clinical Neurophysiology,* 44 (Suppl): 261–272.

Gorodnitsky, I.F., George, J.S. and Rao, B.D. (1995) 'Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm', *Electroencephalograpy and Clinical Neurophysiology,* 95: 231–251.

Grave de Peralta, M.R., Gonzalez Andino, S.L., Morand, S., Michel, C.M. and Landis, T. (2000) 'Imaging the electrical activity of the brain: ELECTRA', *Human Brain Mapping,* 9: 1–12.

Harkins, S.W., McEvoy, T.M. and Scott, M.L. (1979) 'Effects of interstimulus interval on latency of the brainstem auditory evoked potential', *International Journal of Neuroscience*, 10: 7–14.

Hillyard, S.A. and Münte, T.F. (1984) 'Selective attention to color and location: an analysis with event-related brain potentials', *Perception and Psychophysics*, 36: 185–198.

Hjorth, B. (1975) An online transformation of EEG scalp potentials into ortogonal source derivations', *Electroencephalograpy and Clinical Neurophysiology*, 39: 526–530.

Horovitz, S.G. and Gore, J.C. (2004) 'Simultaneous event-related potential and near-infrared spectroscopic studies of semantic processing', *Human Brain Mapping*, 22: 110–115.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V. and Sejnowski, T.J. (2000) 'Removing electroencephalographic artifacts by blind source separation', *Psychophysiology*, 37, 163–178.

Kiebel, S.J., Tallon-Baudry, C. and Friston, K.J. (2005) 'Parametric analysis of oscillatory activity as measured with EEG/MEG', *Human Brain Mapping* 26: 170–177.

Lantz, G., Grave, d. P., Spinelli, L., Seeck, M. and Michel, C.M. (2003) 'Epileptic source localization with high density EEG: how many electrodes are needed?', *Clinical Neurophysiology*, 114: 63–69.

Lins, O.G., Picton, T.W., Berg, P. and Scherg, M. (1993) 'Ocular artifacts in recording EEGs and event-related potentials. II: source dipoles and source components', *Brain Topography*, 6: 65–78.

Liu, H., Gao, X., Schimpf, P.H., Yang, F. and Gao, S. (2004) 'A recursive algorithm for the three-dimensional imaging of brain electric activity: shrinking LORETA-FOCUSS', *IEEE Transactions of Biomedical Engineering* 51: 1794–1802.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T. and Oeltermann, A. (2001) 'Neurophysiological investigation of the basis of the fMRI signal', *Nature*, 412: 150–157.

Luck, S. J. (2005) *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E. and Sejnowsky, T.J. (2002) 'Dynamic brain sources of visual evoked responses', *Science*, 295: 690–694.

Maris, E. and Oostenveld, R. (2007) 'Nonparametric statistical testing of EEG- and MEG-data', *Journal of Neuroscience Methods*, 164: 177–190.

Matzke, M., Mai, H., Nager, W., Russeler, J. and Münte, T. (2002) 'The costs of freedom: an ERP – study of non-canonical sentences', *Clinical Neurophysiology*, 113: 844–852.

McCarthy, G. and Wood, C.C. (1985) 'Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models', *Electroencephalograpy and Clinical Neurophysiology*, 62: 203–208.

Miller, J., Patterson, T. and Ulrich, R. (1998) 'Jackknife-based method for measuring LRP onset latency differences', *Psychophysiology*, 35: 99–115.

Monchi, O., Ko, J.H. and Strafella, A.P. (2006) 'Striatal dopamine release during performance of executive functions: a [(11)C] raclopride PET study', *Neuroimage*, 33: 907–912.

Nunez, P.L. and Srinivasan, R. (2005) *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford: Oxford University Press.

Obrig, H. and Villringer, A. (2003) 'Beyond the visible – imaging the human brain with light', *Journal of Cerebral Blood Flow Metabolism* 23: 1–18.

Osman, A. (1998) Brainwaves and mental processes: electrical evidence of attention, perception and intention', in Scarborough, D. and Sternberg, S. (eds.), *Methods, Models, and Conceptual Issues. An Invitation to Cognitive Science.* Cambridge, MA: MIT Press. pp. 865–915.

Pascual-Marqui, R.D. (2002) 'Standardized low resolution brain electromagnetic tomography (sLORETA): technical details', *Methods and Findings in Experimental and Clinical Pharmacology*, 24: 5–12.

Pascual-Marqui, R.D., Michel, C.M. and Lehmann, D. (1994) 'Low-resolution electromagnetic tomography – a new method for localizing electrical activity in the brain', *International Journal of Psychophysiology*, 18: 49–65.

Perrin, F., Pernier, J., Bertrand, O. and Echallier, J.F. (1989) 'Spherical splines for scalp potential and current density mapping', *Electroencephalograpy and Clinical Neurophysiology*, 72: 184–187.

Pfurtscheller, G. and Aranibar, A. (1977) 'Event-related cortical desynchronization detected by power measurements of scalp EEG', *Electroencephalography and Clinical Neurophysiology*, 42: 817–826.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Jr. Miller, G.A., Rifter, W., Ruchkin, D.S., Rugg, M.D., and Taylor, M.J. (2000) 'Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria' *Psychophysiology*, 37: 127–152.

Rollnik, J.D., Schroder, C., Rodriguez-Fornells, A., Kurzbuch, A.R., Dauper, J., Moller, J. and Münte, T.E. (2004) 'Functional lesions and human action monitoring: combining repetitive transcranial magnetic stimulation and event-related brain potentials', *Clinical Neurophysiology*, 115: 145–153.

Ruff, C.C., Bestmann, S., Blankenburg, F., Bjoertomt, O., Josephs, O., Weiskopf, N. et al. (2008) 'Distinct causal influences of parietal versus frontal areas on human visual cortex: evidence from concurrent TMS fMRI', *Cerebral Cortex* , 18: 817–827.

Rugg, M.D. and Coles, M.G. (1996) 'The ERP and cognitive psychology: conceptual issues', in Rugg, M.D. and Coles, M.G.(eds.), *Electrophysiology of Mind. Event-related Brain Potentials and Cognition.* Oxford: Oxford University Press. pp. 27–38.

Scherg, M. and von Cramon, D. (1985) '2. Bilateral sources of the late AEP as identified by a spatio-temporal dipole model', *Electroencephalography and Clinical Neurophysiology*, 62: 32–44.

Sternberg, S. (1998) 'Discovering mental processing stages: the method of additive factors.', in Scarborough, D. and Sternberg, S. (eds.), *Methods, Models, and Conceptual Issues. An Invitation to Cognitive Science.* Cambridge, MA: MIT Press. pp. 703–863.

Tallon-Baudry, C., Bertrand, O., Delpuech, C. and Pernier, J. (1996) 'Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human', *Journal of Neuroscience*, 16: 4240–4249.

Urbach, T.P. and Kutas, M. (2002) 'The intractability of scaling scalp distributions to infer neuroelectric sources', *Psychophysiology*, 39: 791–808.

Urbach, T.P. and Kutas, M. (2006) 'Interpreting event-related brain potential (ERP) distributions: implications of baseline potentials and variability with application to amplitude normalization by vector scaling', *Biology and Psychology*, 72: 333–343.

Van Veen, B.D., van Drongelen, W., Yuchtman, M. and Suzuki, A. (1997) 'Localization of brain electrical activity via linearly constrained minimum variance spatial filtering', *IEEE Transactions of Biomedical Engineering*, 44: 867–880.

Verleger, R., Gasser, T. and Mocks, J. (1982) 'Correction of EOG artifacts in event-related potentials of the EEG: aspects of reliability and validity', *Psychophysiology*, 19: 472–480.

Wastell, D.G. (1977) 'Statistical detection of individual evoked responses: an evaluation of Woody's adaptive filter', *Electroencephalography and Clinical Neurophysiology*, 42: 835–839.

Weyerts, H., Penke, M., Dohrn, U., Clahsen, H. and Münte, T. F. (1997) 'Brain potentials indicate differences between regular and irregular German plurals', *Neuroreport,* 8: 957–962.

Yeung, N., Bogacz, R., Holroyd, C.B. and Cohen, J.D. (2004) 'Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods', *Psychophysiology*, 41: 822–832.