

An introduction to the measurement of auditory event-related potentials (ERPs)

Remijn, Gerard Bastiaan

Department of Human Science, Faculty of Design, Kyushu University | Research Center for Applied Perceptual Science, Kyushu University

Hasuo, Emi

Faculty of Medical Sciences, Kyushu University | Japan Society for the Promotion of Science

Fujihira, Haruna

Graduate School of Design, Kyushu University

Morimoto, Satoshi

Graduate School of Design, Kyushu University

<https://hdl.handle.net/2324/1912788>

出版情報 : Acoustical Science and Technology. 35 (5), pp.229-242, 2014-01-01. The Acoustical Society of Japan

バージョン :

権利関係 :

TUTORIAL

An introduction to the measurement of auditory event-related potentials (ERPs)

Gerard B. Remijn^{1,*}, Emi Hasuo^{2,†}, Haruna Fujihira^{3,‡} and Satoshi Morimoto^{3,§}

¹*Department of Human Science, Faculty of Design, Kyushu University/
Research Center for Applied Perceptual Science, Kyushu University,
Shiobaru 4-9-1, Minami-ku, Fukuoka, 815-8545 Japan*

²*Faculty of Medical Sciences, Kyushu University/Japan Society for the Promotion of Science,
Maidashi 3-1-1, Higashi-ku, Fukuoka, 812-8582 Japan*

³*Graduate School of Design, Kyushu University,
Shiobaru 4-9-1, Minami-ku, Fukuoka, 815-8545 Japan*

Abstract: In 1939, Pauline Davis reported the first study on event-related potentials (ERPs) performed on awake humans. ERPs are time-locked brain potentials that occur in response to cognitive, motor or perceptual *events*. The events used by Davis were sounds, and in the decades that followed her landmark study ERP research significantly contributed to the knowledge of auditory perception and neurophysiology we have today. ERPs are very well suited to study neural responses to sound stimuli, since the researcher can monitor the brain's registration of sound edges and spectral changes in sound on a millisecond-by-millisecond basis. In this overview we will introduce basic concepts of auditory ERP research. The overview includes descriptions of typical ERP components, experimental paradigms, sound stimuli, research methodology, and ways to analyze data.

Keywords: Event-related potentials, Auditory stimuli, Sound, Speech

PACS number: 43.64.Ri [doi:10.1250/ast.35.229]

1. AUDITORY ERPs

1.1. Introduction

Following work on rabbits and monkeys in the late 19th century, in 1924 Hans Berger succeeded in recording the first human *electroencephalogram* (EEG): a recording of electrical activity of the brain from the surface of the head. By placing electrodes over the scalp and amplifying the output voltage, he managed to plot voltage changes in the human brain during different states of consciousness, such as sleep, wakefulness, attentional engagement, and coma [1,2]. A year after Berger's publication of his findings in 1938, Pauline Davis, in collaboration with her husband Hallowell, reported the first study on *event-related potentials* (ERPs) performed on awake humans in 1939 [3]. In the following decades, auditory ERP research not only added to the basic knowledge about sound processing in the healthy brain. It also greatly improved our understanding of the neurophysiological underpinnings of

various clinical disorders. Nowadays, for example, auditory ERPs are being used to study the neurodevelopmental status of persons with language development problems [4], Alzheimer's disease [5], and autism [6].

Below we will give an introduction to the measurement of ERPs in response to sound ("*auditory ERPs*"). ERP measurement is very well suited to study the brain's response to sound stimuli. Most natural sounds in the world vary in spectral content over time and the brain's registration of these changes can be monitored through ERP measurements on a millisecond-by-millisecond basis. Furthermore, auditory ERPs can be obtained in response to a wide variety of sounds, and experimental paradigms can be employed with either passive listening or active listening, i.e., combined with a behavioral response to the sound or a part of the sound. Finally, auditory ERPs can give valuable information about certain stages of auditory processing, either with or without attentional engagement of the listener to the event.

We will start this overview with a description of ERPs from a physiological viewpoint and some basic, general ERP terminology. We will describe a number of so-called ERP components and paradigms that have been identified

*e-mail: remijn@design.kyushu-u.ac.jp

†e-mail: hasuo@neurophy.med.kyushu-u.ac.jp

‡e-mail: h.fujihira.228@s.kyushu-u.ac.jp

§e-mail: satoshi.morimoto.225@s.kyushu-u.ac.jp

and used over the years, and discuss what sensory and/or cognitive stages of sound processing they reflect. This is followed by a description of the basic ERP measurement environment, with focus on typical equipment and participant instructions. We will finish by mentioning typical analysis methods. Please bear in mind that this overview deals with only the most elementary issues of auditory ERP research. Where possible we will include suitable references and refer the reader to review papers on particular topics for more advanced information. We sincerely hope the review supplies enough basic ERP knowledge to raise the reader's interest in doing research in the field.

1.2. Basic Neurophysiology and Terminology

ERPs recorded from the human scalp reflect the electrical activity, produced by sinks and sources, of populations of pyramidal cells. These cells form the grey matter of the surface of the cerebral hemispheres, i.e. the cerebral cortex. ERPs are extracted from the ongoing brain activity, as reflected in the EEG. A major issue for researchers in the early ERP-days was that the EEG is a raw mixture of ongoing brain activity that is generated by numerous neural sources. So, how did they extract the ERP from the mixture? The answer is that they did not—most of the early research is based on observations from single trials in which signal-to-noise ratios were favorable. In the 1950s, the study of ERPs became increasingly more systematic. Dawson [7], for example, started to extract a larger number of ERPs from the EEG; simply by increasing the number of event recordings he obtained an increased signal-to-noise ratio. The advent of the computer made ERP isolation even easier. The first study with computer-averaged ERPs was allegedly published in 1962 by Galambos and Sheatz [8]. Nowadays, ERP research is routinely performed with the aid of computers and multiple repetitions of the same event. Signal-to-noise ratios can be controlled far better than in the early days, thanks to vast technological improvements, but we still need to deal with all kinds of 'noise,' such as external noise and internal noise generated by the participant from whose scalp we are recording (see Sect. 2.3.3). Therefore, in every ERP experiment, repetition is the key. The necessary total amount of same-event repetitions varies with the experimental paradigm and the ERP component of interest, but the general rule is that more trials will result in cleaner averaged waveforms.

What does all this stimulus repetition and averaging lead to? Averaging the brain activity during fixed time-windows locked to the repetitive appearance of an event of interest should eventually result in a waveform that consists of a series of overlapping peaks and troughs. These peaks and troughs are the *ERP components*. In its most

fluid definition, an ERP component is scalp-recorded activity generated by a specific neural or physiological process, which in turn produces a certain polarity, latency, and scalp distribution (for an excellent review see [9]). As for *polarity*, ERP components show voltage deflections (expressed in μV) that are either positive or negative. Animal studies have suggested that the polarity of ERPs is related to either excitation or inhibition of cells, yet with a complex neural generation process (see [10] for a review). Simply put, neuronal discharges/firing in thalamo-cortical cells generally result in negative ERP components, while cellular inhibition underlies positive potentials [11]. With some exceptions, ERP components are indicated with a letter that indicates their polarity—"P" for positive and "N" for negative. This letter is typically followed by a number that indicates the *latency* of occurrence. ERP components are time-locked to a stimulus or a task. A researcher could be, for example, interested in the onset of a sound or a frequency change within a sound. He/she could then time-lock the start of ERP extraction to that particular event. Typical ERP components in response to the start of a sound are the N100 and the P200. The N100 is a negative deflection with an average latency that can be identified in the ERP waveform at about 100 ms following event onset. The P200 is a positive deflection with an average latency of about 200 ms after event onset. We will discuss the N100-P200 complex and a few other components below in Sect. 1.3.

The *scalp distribution* of an ERP component refers to the locations over the scalp where the component most prominently occurs during the processing of an event. The brain can be roughly divided into the two hemispheres, each with frontal, temporal, parietal, and occipital areas that are covered by different EEG electrodes. Most ERP components, including those in response to sound, have a wide scalp distribution. For example, one of the most-studied functional ERP components, the so-called mismatch negativity (MMN; see Sect. 1.3) is negative in voltage and occurs over frontal and central scalp areas with reversal into positivity below the Sylvian fissure. Neural generators of the MMN in response to sound have been localized at the superior temporal gyrus near Heschl's gyrus, an area considered to mediate basic feature processing in auditory cortex. Other neural generators of the MMN have been identified as well, reflecting the fact that what we observe at the scalp is actually a mixture of activity from many different brain areas. Of interest in general is the change in scalp distribution over time and the difference between scalp distributions in response to different stimuli or tasks. Even though we often cannot know the exact neural sources of ERP components, we can infer that differences in scalp distribution indicate different cortical processing activity.

Next to descriptions based on properties such as polarity, latency, and scalp distribution, ERP components are conventionally divided into exogenous and endogenous ERP components [12]. The exogenous components of the ERP trace are said to occur within the first 100 ms after event onset. Here the transmission of incoming information from the peripheral sensory system to the cortices is reflected in the ERP deflections. The characteristics of these ‘sensory’ ERPs reflect the physical properties of sensory stimuli. Endogenous ERP components, on the other hand, reflect the psychological or cognitive significance of an event or the higher-order processing of that event. These ‘psychological’ ERPs appear after approximately 200 ms and their properties vary with the processing demands of the stimulus or task, such as the engagement of attention. Additional research has shown that the distinction between exogenous and endogenous ERP components is certainly not a strict one. Attention, for example, may have an influence on the properties of exogenous components in a top-down way. Nowadays, descriptions of ERP components mainly focus on the cognitive processes they reflect (e.g., [13]).

Below we will give a brief overview of some of the most studied ERP components in relation to auditory stimuli and processing. We will concentrate on the cortical responses in the temporal order they appear in the waveform. Subcortical responses are not discussed. Instead, we refer readers who are interested in EEG research on auditory subcortical responses, e.g., brainstem responses, to the review by Stone and colleagues [14]. Auditory brainstem responses consist of a series of small voltage potentials that can be visualized in response to a click-train with a few hundred repetitions or more. Often measured in clinical settings, for example to test hearing loss and nervous system disorders, auditory brainstem responses are short-latency responses that occur within approximately 10 ms from click onset.

1.3. A Selection of ERP Components and Paradigms

The P50 component. The P50 is a positive deflection that occurs approximately 50 ms after sound onset. Pratt and colleagues (2008) called it “the earliest, the smallest in amplitude, the most variable and consequently the least studied of the auditory ERPs.” Although its amplitude is relatively small, it robustly follows the onset of all kinds of sounds, including short clicks or noise bursts. It has been argued that the P50 reflects pre-attentive arousal due to the appearance of a new event in the auditory scene. Since off-responses to sound do not comprise this “warning” element of sound, the P50 is weak, or often absent, after sound offset [15]. The P50 is known to have a reduced response to a second stimulus, as compared to that to a first stimulus, if the stimulus pair is separated by about 500 ms [16]. The

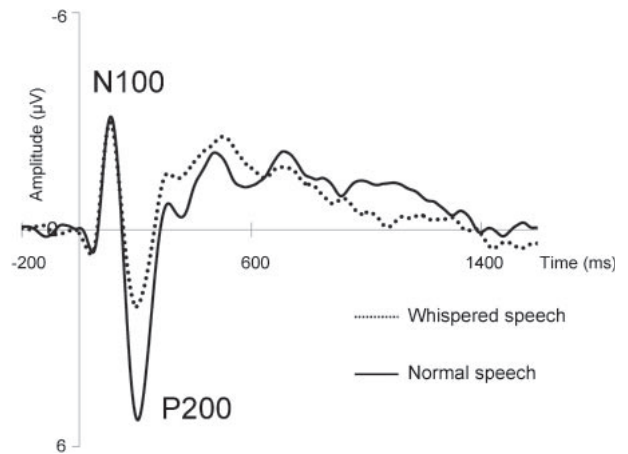


Fig. 1 N100-P200 ERP components in response to normal and whispered speech. The P200 amplitude difference is caused by the intensity difference between the speech stimuli. Note that amplitude is depicted as “negative-up.”

diminished response to the second stimulus is caused by inhibition by the first stimulus. This finding has been exploited in schizophrenia research, which showed that schizophrenia patients show less second-response suppression than healthy controls (e.g., [17]). Although the P50 is thought to reflect pre-attentive processing, the schizophrenia research thus implies that even the very earliest ‘sensory’ ERP component reflects the psychological state of the listener (e.g., [18]).

The N100 component. The N100 is a relatively large, negative deflection that reliably occurs approximately 100 ms after an abrupt change in the auditory environment [19]. The N100 (Fig. 1) has been described as reflecting neural processes that are sensitive to stimulus features such as the “suddenness” of the sound’s onset and offset, i.e., the sound intensity at the edges (e.g., [20]). Hillyard and colleagues in 1973 [21] were among the first to investigate the N100 systematically. To this end they used the shadowing task in a *dichotic listening* paradigm, well-known in the field of auditory cognition and selective attention, in which listeners are asked to verbally repeat information entering one ear, while ignoring different information presented to the other ear. The authors found that the ‘sensory’ N100 component too was subject to attention (see also [22]).

The P200 component. The P200 (Fig. 1) is a positive deflection occurring approximately 200 ms after sound onset. Similar to the N100, the P200 is a salient deflection that reliably occurs across individuals. The P200 amplitude varies with the physical characteristics of a sound, such as its intensity [23] and frequency [24]. It also varies with the acoustic properties of speech sounds, for example those signaling prosodic content [25]. Figure 1 shows the

N100-P200 complex in response to normal and whispered speech [26].

Mismatch negativity (MMN). MMN is a negative deflection in the EEG waveform that typically occurs in between 150–250 ms after stimulus onset. It was first described by Näätänen and colleagues [27] who used a so-called *oddball paradigm*. In this paradigm, MMN occurs in response to an infrequent stimulus (the *deviant*) that occurs within a stream of frequent stimuli (the *standard*). MMN is considered to reflect early sensory memory processing involved in matching the incoming stimulus with that of previously processed stimuli. The occurrence of MMN does not require attentional engagement to the sounds. Whereas the N100 amplitude, and that of other components, is larger when a participant pays attention to the stimuli [28], MMN can be obtained even when participants ignore the stimuli [29,30] or when they are asleep [31]. In order to obtain MMN during sleep, however, the difference between the deviant and standard must be prominent.

With auditory stimuli, MMN can be elicited by an oddball stimulus that differs from the ongoing stream of stimuli with regard to, for example, intensity, frequency, and duration. Slightly more complex stimulus configurations also elicit MMN, for example when a sequence of tones that consistently increases in frequency is interrupted by a deviant with the same frequency as the former tone [32]. In speech, MMN research has been often used to obtain neural correlates of *voice-onset time (VOT)* discrimination in listeners (e.g., [33]). MMN is obtained through calculation of a *difference wave* between the waveforms induced by an equal number of frequent and oddball stimuli. The resulting waveform tends to peak around 200 ms after stimulus onset, but latency and amplitude vary with factors such as *inter-stimulus interval (ISI)*, the ratio of occurrence of frequent and deviant stimuli, the difference in stimulus characteristics between frequent and deviant stimuli, and so forth.

The N200 component. It is interesting to note that the auditory MMN is also called the N2a, a subcomponent of the N200. The N200 is a negative deflection typically evoked in between 180–325 ms following the presentation of a deviant stimulus. The N2b component of the N200 is a component that follows the N2a. Contrary to the N2a, the N2b is evoked when the participant pays attention to the stimuli and it often occurs in tandem with the P300 (see below). The N2c occurs in classification tasks. For further reading on MMN and the N200 and its components we refer to Näätänen and colleagues (on MMN [34]) and Pritchard and colleagues (on N200; [35]).

Object-related negativity (ORN). ORN has been employed in studies on *concurrent sound processing*. ORN is a negativity that peaks after approximately 140 ms in

response to a mistuned tonal component in a complex sound. Like MMN, ORN can be observed after subtraction of the waveforms in response to a complex sound without and with the mistuned tone [36]. ORN is said to reflect pre-attentive processes of perceptual organization of multiple auditory events. When asked to make a perceptual judgment about the incoming sounds, e.g., whether the sounds occurred simultaneously or not, often a positive deflection occurring approximately 400 ms after stimulus onset can be observed as well. Whereas ORN is considered to reflect pre-attentive processing, this *P400 component* is thought to reflect attentive processes involved in object evaluation.

The P300 component. The P300, a positive deflection occurring approximately after 300 ms in the waveform, is well-studied for its reflection of (in)voluntary *attentional engagement* to a stimulus. As with MMN, the P300 can also be elicited with an unexpected stimulus in an oddball paradigm. Following the first report of the P300 by Sutton and colleagues in 1965 [37], a wealth of research has been performed with the P300 and nowadays researchers target one of its subcomponents, often the P3a or the P3b.

The P3a has a more frontally-oriented scalp distribution than the P3b. The P3a is supposed to reflect *involuntary attention* to a novel stimulus in an oddball paradigm in which behavioral responses to the stimuli are not required. The P3b component of the P300, by contrast, is strongly associated with *voluntary attention* to the stimuli in an oddball paradigm that does require behavioral responses. The peak amplitude of the P3b varies with the relative amount of deviant stimuli, as compared to the number of standard stimuli, while stimulus relevance and the availability of attentional resources affect the amplitude as well [38]. The latter in particular has raised the interest of clinical researchers involved in the study of attention-related deficits, such as attention deficit hyperactivity disorder (ADHD). It also appears in ‘*go/no-go paradigms*’ in which participants are asked to respond to one stimulus (the go stimulus), and suppress their responses to another stimulus (the no-go stimulus). Participants, for example, perform a task in which reaction time is measured with go/no-go stimuli that differ in intensity.

The N400. The N400 is a negative deflection occurring approximately 400 ms after the presentation of visually-presented or spoken words. It was first reported by Kutas and Hillyard in 1980 [39] and has since played an important role in the study of *semantic processing* of speech. The N400 has been reported in studies that used a wide variety of paradigms, such the shadowing task in a dichotic listening experiment and the *sentence-completion paradigm*. In spite of numerous studies on the N400, its exact functional meaning still eludes the research community as reviewed by the researchers who first discovered and reported the N400 [40].

Contingent negative variation. Contingent Negative Variation is generally evoked in a classic paradigm in which a *warning stimulus* is followed by a so-called *imperative stimulus*, which urges a behavioral or a mental response from the participant. First described by Grey Walter *et al.* in 1964 [41], CNV is a sustained negativity that can be observed in between 250–450 ms after the warning stimulus, during the interval in between the warning stimulus and the imperative. CNV is mainly considered to reflect response preparation and it varies with stimulus features, such as intensity, probability, relevance, pitch, and duration (e.g., [42]) as well as attentional engagement [43].

Please bear in mind that the list of components and paradigms mentioned above is far from exhaustive. We have mentioned only the most common ERP components studied with auditory stimuli, known to appear robustly and reliably given the right paradigm is used. Subcomponents, variations in paradigms, and less well-known components have been reported in the wealth of literature on ERPs. Among the components we did not discuss, some are predominantly exploited in visual research, such as the N170 (mostly concerned with face processing) and the P600 (association with memory/recall of visually presented materials). Other components may be difficult to consider for those early in their ERP-research career, such as Left Anterior Negativity (LAN) and the Late Positive Potential (LPP). The components and paradigms we have mentioned above are all meant to eventually clarify cortical processing, either at the sensory or higher-order, cognitive level. ERP research should ideally proceed from the cortical processes in the brain that one wants to study. After the research question has been formed, the focus should be on finding the suitable paradigm and ERP component(s) that can explain or clarify the cortical process of interest.

2. THE AUDITORY ERP EXPERIMENT

2.1. ERP Recording with EEG Equipment

2.1.1. Basic EEG recording setup

The EEG recording system usually consists of electrodes, an EEG amplifier, an analog filter, an analog-to-digital (A/D) converter, and a computer with EEG acquisition software [44]. *Electrodes* are placed on the participant's scalp and face, and are used to pick up the electric signals generated from the brain. The *amplifier* brings the microvolt signals captured by the electrodes to some range in which the signals can be digitized accurately. *Analog filtering* is performed at the same time as the amplification to cut the very high frequencies (to prevent aliasing in the A/D converter), and the very low frequencies (to prevent blocking the converter by slow changes in the baseline; [44,45]). The *A/D converter*

converts the signals from analog to digital form, and the digitized data are recorded in the computer. In many cases, the whole recording system comes in a package from the manufacturer, so there is no need to worry about the waveform acquisition of each device. It is important, though, to consider and specify the data-gathering settings: the type of electrodes, inter-electrode impedance, the location of recording electrodes on the scalp, the gain or resolution, and the filtering characteristics of the recording system [44], which we will discuss in following sections.

When conducting experiments to measure auditory ERPs, equipment to generate and to present the auditory event (e.g., a stimulus presentation computer and headphones or speakers) will be needed in addition to the EEG equipment. When using *headphones*, shielded headphones might be considered, because most audio headphones can cause electrical artifacts that will be picked up by the electrodes. For later averaging of the EEG signals for each stimulus condition (or for each response by the participant), a *trigger signal* needs to be sent with the stimulus event, or with the response to that event. We recommend the work of Picton and colleagues [44] and Luck [45] for more detailed information and helpful tips about setting up the recording environment.

2.1.2. Active, reference, and ground electrodes

A typical EEG system for gathering ERP data uses a so-called differential amplifier, and at least three electrodes are required: an active electrode (A), a reference electrode (R), and a ground electrode (G). The *active electrode* is placed at the desired scalp location (often at sites over which activation is expected; see Sect. 2.3.1). Traditionally, the *reference electrode* (R) is placed at the earlobe, the mastoid (the bony part behind the earlobe), the nose, or elsewhere on the scalp. The reference can even be computed off-line and the ideal reference in the field has changed over the years. Most importantly, the reference should not affect the characteristics of the ERP component of interest, such as its latency, amplitude and scalp distribution. The *ground electrode* (G) is placed at some convenient location on the participant's head or body, and again places such as the mastoid or the forehead can be considered. Given the number of possible AGR-electrode configurations, it is recommendable to get acquainted with the trends in the latest literature.

The *differential amplifier* of the EEG system amplifies the difference between the AG voltage and the RG voltage, i.e., the voltage is first measured between A and G, and between R and G, and then the output will be the difference between these voltages: $(A - G) - (R - G) = A - R$. The ground (G) thus does not exist in the output result. This means that ambient electrical activity picked up by the ground circuit (i.e., noise), which is the same for AG and

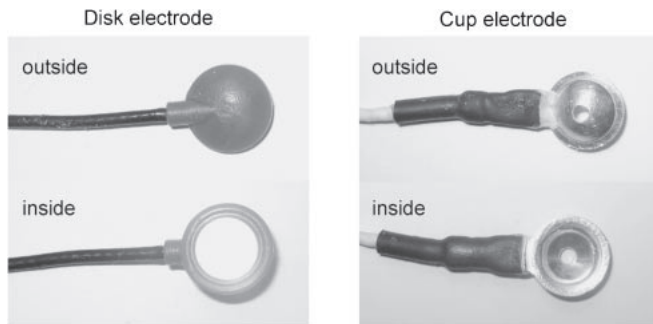


Fig. 2 Disk electrode (left) and cup electrode (right) used in ERP experiments.

RG, will be eliminated by subtraction [45]. Note that the noise will not be subtracted away perfectly if the AG and RG signals are not treated equivalently, and this is one reason to keep the electrode impedances low (see Sect. 2.3.2). Moreover, it is important to keep in mind that an ERP waveform reflects the difference in activity between the active and reference sites, not just the activity at the active electrode. ERP waveform characteristics thus can look different with different references. Comparing ERP data and experiments is therefore not always straightforward if different AGR configurations have been used.

2.1.3. Different types of electrodes

The electrodes are usually metal disks or cups, attached to a wire (Fig. 2), which can be placed upon the scalp with a *conductive paste* or *gel*. The choice of metal is important (see [46]). Silver electrodes covered with a coating of silver-chloride (typically called Ag/AgCl electrodes) are said to be able to accurately record very slow potential changes (e.g., [47]). Other materials such as gold or tin may be used for fast-changing potentials [44].

When using large numbers of electrodes an elastic cap (Figs. 3a–3c) or net (Figs. 3d–3f) can be helpful. The electrode types may differ between cap and net systems, and procedures for electrode attachment onto the scalp and impedance reduction again may differ depending on the type of electrode. For example, the *cap electrodes* usually have a hole in the middle (Figs. 3b, 3c), and the impedance is reduced by inserting a syringe through the hole to abrade the underlying skin and to inject a conducting gel.

Net electrodes are usually embedded in sponges (Fig. 3f; [48]), which need to be soaked in a KCl solution before being fitted on the scalp. With this system, abrasion of the skin is not necessary, so it is possible to attach large numbers of electrodes quickly (the system offers up to 256 channels). There are many different electrodes available from different manufacturers, and it is important to choose the right type considering the purpose of the study [45].

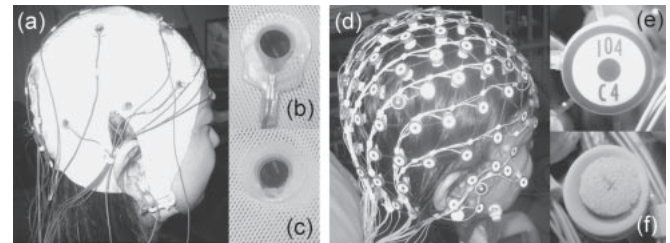


Fig. 3 Example of an electrode cap (Eldith, NeuroPrax) applied to a participant (a), with a typical cap electrode shown from the outside (b) and inside (c). Example of a sensor net (Electrical Geodesics, Inc., Geodesics Sensor Net) applied to a participant (d), with a typical net electrode shown from the outside (e) and the inside (f).

2.2. Sound Stimuli and Preparation

Since the EEG waveform can reflect sensory and cognitive processing on a very acute time scale, *sound stimuli* have been used extensively in ERP research over the years. From the research mentioned in the description of some common ERP components (Sect. 1.3), we have seen that sounds ranging from simple sounds, like clicks or tone bursts (e.g., [49,50]), to complex sounds such as music (e.g., [51,52]) or speech (e.g., [53]) can be employed. Sound stimuli can also be presented in multi-modal research, with stimuli of other sensory modalities (e.g., [54]). Basically, the same auditory stimuli as in behavioral experiments can be used in ERP studies, so brain mechanisms underlying perceptual phenomena, such as *auditory illusions*, can be investigated with ERPs as well. For example, the well-known octave illusion [55,56] and the filled-duration illusion [42] have been investigated with ERPs. Neural correlates of temporal assimilation [57] have been clarified [58], and the classic phenomenon of binaural beats [59] was studied with ERPs almost 60 years later [60].

2.2.1. Tips for sound stimuli preparation

When preparing the sound stimuli for an ERP experiment, the same precautions as when preparing the stimuli for auditory psychological experiments should be taken in order to ensure reliability of the sound presentation. For example, it is recommended to include a *rise and a fall portion* at the beginning and the end of a sound to avoid abrupt increase/decrease in sound intensity that could cause *spectral splatter*; the *presentation level* and the *frequency range* of the sound stimuli should be in a range which can be presented without distortion by the sound presentation system; noise stimuli should not be reused (it should be generated for each presentation); and it is always recommendable to measure, record, and analyze the output from the headphone/speakers to check whether the sounds are really presented as intended by the experimenter (e.g., [61,62]).

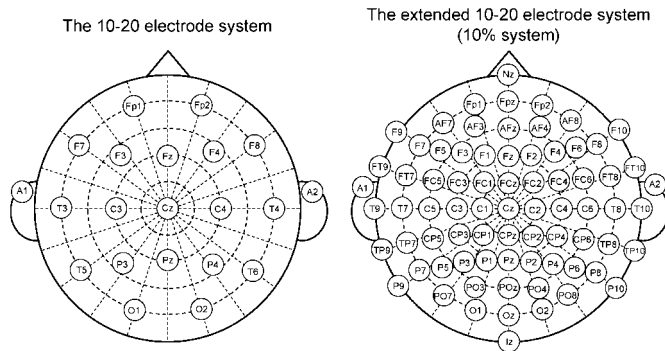


Fig. 4 The international 10-20 system (left) and the extended 10-20 system (right) for electrode placement.

In addition to these points, for ERP experiments it is important to optimize sound characteristics to observe the desired ERP component clearly. Increasing the *intensity* of the auditory stimulus can increase the peak amplitude of auditory evoked potentials (e.g., [49,63]); or as reflected by the difference between normal and whispered speech in Fig. 1). It is also important to note that the amplitude of the evoked potentials can be influenced by the *inter-stimulus interval* with other sounds. Generally, increasing the ISI increases the peak amplitudes of N100 and P200 to the sound (e.g., [64,65]), while ISIs in the range below 1 s decrease the amplitude of change-N1, i.e., a subcomponent of N100 that is elicited by a change within an auditory stimulus [66].

2.3. Experimental Procedure

2.3.1. Attachment of electrodes

The electrodes or electrode caps as described in Sect. 2.1.3 should be fitted on the standard electrode positions, either following the *international 10-20 electrode system*, the extended 10-20 electrode system (10% system), or the 10-5 electrode system (Fig. 4). The 10-20 electrode system is a traditional system that contains 21 *electrode locations*. It was introduced in 1958 by Jasper [67], and it is still used today. In the 10-20 system, the scalp surface is divided into percentages proceeding from anatomical landmarks. Two of these landmarks are the *nasion* (the depressed area between the eyes) and the *inion*, located at the back of the head where the skull is at its lowest point (typically a bump straight above from the neck). If we measure the distance between the nasion and the inion, as well as the distance between the ears from the locations just in front of the earholes, we can determine scalp location Cz. This is the location on the scalp where our center electrode will be placed. From here on, following the 10-20 system, electrodes can be placed over frontal, parietal, and temporal areas indicated by the letter F, P, and T, respectively. Besides a letter, scalp/electrode locations have an odd number for left-hemispheric loca-

tions and an even number for right-hemispheric locations. When using sound, we are often interested in so-called *midline locations* (Fz, Cz, and Pz), as well as locations T3 and T4, among others, because these locations cover the auditory cortices.

The most recent multi-channel EEG hardware systems enable the researcher to use a far larger number of channels than the 21 locations indicated in the 10-20 electrode system. For proper electrode/cap fitting, the 10-20 system has therefore been extended, with the 10% system [68–70] and the 10-5 electrode system [71]. The 10-5 electrode system contains more than 300 electrode locations and can be considered when an electrode net is used (cf. Fig. 3d). In the case of *electrode cap fitting*, the researcher must take care of the following. First, each electrode must be placed over the correct designated electrode location, and one must beware of the possibility that the electrode cap can shift over the scalp during the experiment. In order to avoid this, the cap should be attached to a body harness that must be used correctly. Next, care must be taken not to use too much gel or paste to connect the electrodes to the scalp. When the number of electrodes is large, the space between the electrodes is small and excessive gel can spread in between two or more electrode locations. As a result, similar changes in waveform potentials will be recorded from different electrode positions.

2.3.2. Controlling electrode impedance

When the electrodes are fitted onto the scalp, it is important to check *electrode impedance*. Impedance is measured by passing very low currents through the electrodes. According to the ERP publication guidelines [44], the measurement of electrode impedance tells the experimenter four things: how accurately the amplifier will record the potentials, the liability of the electrodes to pick up electromagnetic artifacts (‘noise’), the ability of the differential amplifiers to reject common-mode signals, and the intactness of the skin underlying the electrode. Electrode impedance should generally be reduced to less than 5 k Ω . In order to reduce the electrode impedance, the scalp and skin must be cleansed effectively from sweat, sebum and scurf. The most common way to do this is by cleaning the scalp and skin with cotton saturated with ethanol. Further use of abrasive compound is also effective. However, since the scalp and skin can be sensitive, the experimenter must be careful not to clean too rigorously in order to avoid the risk of infection and pain. First and foremost, it is necessary to check with the participant whether he/she has no allergic reaction to certain substances, ethanol for example, before the experiment starts. During the experiment it is important to check whether the participant feels no pain or is otherwise uncomfortable, and to make sure that ethanol or gel does not reach the participant’s eyes or clothes.

2.3.3. Controlling measurement artifacts

ERP measurements are subject to many different kinds of artifacts generated from sources outside the brain. Since the responses from inside the brain are the single center of focus, *artifact contamination* must be avoided or reduced as much as possible. Some artifacts arise from the body of the participant. The first types of artifacts that must be dealt with are *ocular artifacts*, which occur due to eye movements and blinks. The eye forms an electric dipole because the cornea is positively charged, while the retina is negatively charged. Eye movements and blinks therefore cause a change in the electric fields that surround the eyes, which distort the electric fields over the scalp and, hence, the EEG measurement [72,73]. There are some ways to reduce ocular artifacts. Having participants stare at one fixed point in the experimental room/booth can generally reduce eye movements. Asking participants to open their eyes slightly (to squint) usually reduces heavy eye blinking. If the paradigm allows it, the total amount of ocular artifacts can be limited simply by having short measurement sessions with breaks in between, in which the participant can rest his/her eyes. Besides this, it is common to use a rejection procedure to omit trials contaminated by heavy eye movements from further analysis. *Electro-oculogram* (EOG) recording, often a standard procedure available in the EEG equipment and software, shows the voltage from a series of electrodes close the eyes. EOG can be used to remove trials in which the EOG voltage exceeds a preset criterion [72]. The criterion varies from study to study, but tends to be in a range of 75 to 120 μV [74–76]. It is nevertheless necessary, however, to monitor small-voltage EOGs that passed the preset criterion, since these can still affect EEG results. Another method to remove ocular artifacts is the *multiple source eye correction* (MSEC) method [77], in which the eye activity is estimated (and later corrected) according to both the spatial distribution of eye activity and a model of brain activity. The MSEC method is now widely used (e.g., [78–80]).

In addition to ocular artifacts, other artifacts generated by the participant's body are body movements, myogenic potentials, and sweating. *Body movements* cause fluctuations in the EEG baseline and it is therefore important to instruct the participant to remain in a fixed bodily position during measurements. The *myogenic potential* is generated by muscle contraction, and is commonly a cluster spike with a frequency of some tens of Hz. Actions such as tooth gnashing, swallowing, facial movements, and the straining of muscles in the head or body result in a myogenic potential. Again, the participant should try to control such actions as much as possible. Finally, sweating also causes fluctuations in the EEG baseline. The temperature in the measurement room/booth should thus be kept comfortable.

2.4. Participants

Before starting an auditory ERP experiment, it is important to understand how *participant factors*, such as gender, age, handedness, physical/mental health, and the level of expertise on a certain task, affect ERP components. Regarding *gender*, there are some reports that N100 latencies in females are shorter and have larger amplitudes than in males [81–83]. Support for gender differences in MMN latencies and amplitudes has also been reported [84], but the effect is certainly not robust (e.g., [85,86]).

The effects of *aging* on ERP component characteristics are complex. The responses of adults generally show the N100-P200 complex with N100 at about 100 ms. Toddlers, by contrast, show a large positivity at about 140 ms, followed by a negativity at about 250 ms. The latencies of these components decrease with maturation until the late teenage years (for a review see [87]), and are prolonged in aging adults [88–90]. ERP research with very young children is generally difficult, due to motion artifacts. Some reports have shown that MMN in newborns and preschoolers resembles adult MMN [34,91]. MMN to speech stimuli in 7- to 11-year-old children, however, had a similar latency as adult MMN but a larger amplitude [92]. Elderly participants, by contrast, often show a smaller MMN amplitude than young participants [93]. One reason for this could be the natural decline in the functioning of sensory organs, such as hearing loss. Regardless of age, one of the most essential tasks of the experimenter before starting an auditory ERP experiment is to obtain an *audiogram* from all participants.

As regards *handedness*, it is common practice in brain scientific studies to gather right-handed participants. Left-handed or ambidextrous participants are more prone to display atypical hemispheric specialization, for example for language [94] and music [95]. The (surprisingly) few studies on the topic indeed suggest that handedness can affect ERP data, and hence left-handers are routinely excluded from experiments. Next to the audiogram, before the ERP experiment it is recommendable that the participant performs a *handedness test*, such as the Edinburgh inventory [96].

Given that basic participant factors (gender, age, handedness) and differences in possible task-related factors (e.g., musical training [51,52]) can cause individual differences in the EEG waveform, it is highly recommendable to systematically *record participant information*. Furthermore, it is necessary to match participants for such factors when ERP responses to different tasks or stimuli are measured over different groups. Unless, of course, the researcher is interested in comparing group effects of developmental or (neuro-)physiological differences on ERP components. In that case, ideally the control partic-

ipants should differ from the experimental participants only on the factors being investigated [44].

2.5. Informed Consent and Instructions

Before the start of any (auditory) ERP experiment, obtaining *informed consent* from the participant is mandatory. This is not just to ensure the collaboration of the participant; for any research involving human participants the ERP research proposal must be reviewed and acknowledged by the ethical committee of the university or research institute. Bear in mind that written—not verbal—informed consent is the standard, and publication guidelines for all international, peer-reviewed journals explicitly state that the experimenter must be able to provide informed consent documents from all participants.

Next to approval from the ethics committee and informed consent from the participants, the experimenter must carefully prepare *task instructions*. These are extremely important. Instructions are necessary to limit artifacts such as those discussed in Sect. 2.3.3, but also to get some control over the *internal state* of the participant. Even with a homogenous group of participants as regards gender, age, and other factors, participants differ in their level of arousal and/or attentional engagement towards the experiment. These internal factors strongly affect ERP component characteristics.

We have seen in Sect. 1.2 that attentional engagement modulates ERP components [28], even those considered to reflect pre-attentive or ‘sensory’ processing. Although some auditory ERPs can be obtained during sleep [31], in the vast majority of ERP experiments the participant should stay wide awake and alert. When he/she is executing a behavioral or mental response to an event, attentional engagement is usually secured. However, when the participant is asked to just passively listen to auditory events, controlling the level of alertness is difficult since repetition of same-event trials may soon cause drowsiness. In *passive listening* experiments, it is therefore desirable to give the participant a task to perform (e.g., read a book, solve a puzzle) in order to have some control over what he/she actually does [44]. To keep the participant in an alert state, it may also be useful to provide short breaks among short sessions. ERP measurement sessions tend to be long because a large number of trials is necessary to improve signal-to-noise ratios when averaging the waveforms. Short breaks may help to prevent boredom or sleepiness, of course with some encouraging talk of the experimenter or some refreshments. In any case it is important to ask the participant to sleep well on the day before the experiment.

Next to these attempts to control the participant’s internal state, instructions must be provided to control external bodily noise as well. We have discussed the

harmful effects of eye movements, blinks, body movements, gnashing of teeth and swallowing (see Sect. 2.3.3), and participants should limit these the best they can. However, if the instructions are too strict the participant may be uncomfortable or nervous during measurements, which in turn can cause an increase in the myogenic potential or sweating. Ideally, the instructions must be given in such a way that the participant feels relaxed and comfortable, but not too much, in order to avoid drowsiness or sleep. The preparation of instructions thus is a delicate task indeed. In order to check whether the participant follows the instructions and to observe his/her general well-being, the use of a camera and/or an intercommunication device in the experimental booth may be useful. Of course, he/she should be informed about this. After the experiment, it is recommendable to debrief the participant. *Debriefing* can provide information about how the participant viewed the task and what cognitive strategies he/she used. In any case, it is important to have clear instructions and to check whether the participant understands them well. More detailed information on instructions (and related topics) can be found in the ERP publication guidelines [44].

3. DATA PROCESSING AND ANALYSIS

3.1. ERP Data Processing

Once the EEG data are gathered, an important step in basic waveform processing is to segment the original continuous EEG signal into *epochs*, based on the trigger information associated with the ERP component(s) of interest. For each particular experimental condition, epochs that cover a certain temporal window can be averaged. The epochs should contain a pre-trigger period, or a *baseline*, that is usually longer than 100 ms. The baseline of each epoch can be corrected by subtraction of the mean amplitude over the pre-trigger period. As mentioned, the averaging process improves the *signal-to-noise ratio*, which improves as the number of averaged epochs increases. However, even by averaging a large number of epochs and the application of *baseline correction*, large artifacts cannot be cancelled out completely. Several manipulations have therefore been proposed to remove large artifacts. These can be applied either before or after the epoch segmentation.

Filtering is the most popular method to attenuate unwanted activity in the EEG signal. For ERP studies, both high-pass and low-pass filtering can be applied. *High-pass filtering* is commonly performed over a range of 0.5–1 Hz, while *low-pass filtering* is often done in the range of 30–70 Hz. It needs to be noted that high-pass filtering over short EEG epochs can decrease reliability of the final result. Since high-pass filtering distorts the start and end of the waveform, filtering is basically applied to the contin-

uous waveform before epoch segmentation (for more detail, see the discussion in [44] and [45]). Meanwhile, additional low-pass filtering, with a cutoff frequency of about 20 Hz, is sometimes applied after finishing all the waveform processing. It is typically applied for better visualization of the results, for example in order to make an illustration of amplitude-over-time changes in the averaged waveform that covers the ERP component(s) of interest.

We have mentioned in Sect. 2.3.3 that ocular artifacts can cause large distortions in the EEG waveform. Before starting the averaging (and filtering) process, it is therefore important to reject epochs based on pre-set EOG criteria. For ocular artifacts and other heavy distortions in the waveform, however, one must bear in mind that too many epoch rejections can bias the final outcome and make interpretation difficult [44]. If the artifacts are limited to one or two participants only, one might consider excluding the data from the statistical analyses. A less rigorous method is to try artifact removal through *independent component analysis* (ICA; [97]). ICA is a blind separation technique that decomposes the original signal into statistically independent components [98], proceeding from the assumption that the sources of the waveforms and the artifacts are spatially independent. Simply put, with the help of ICA software the independent components corresponding to artifacts can be identified by eye, and removed. From the remaining components, a new waveform without the artifacts can be reconstructed. To help the EEG experimenter, several automatic ICA methods for artifact correction have been proposed (e.g., [99]).

3.2. ERP Data Analysis

Conventional EEG waveform analysis is performed on the data from each separate electrode location. The most fundamental way to analyze the characteristics of the averaged waveform over a certain time window is first *to label* the troughs and peaks, that is, to identify the ERP components. Following this, one can *measure the amplitude and latency* of each ERP component of interest. As mentioned, the amplitude of a component varies with sensory or cognitive processing, while the latency shows the computational time needed in the brain to perform such processing. If correctly identified and measured for each participant, the experimenter can then statistically compare amplitudes and latencies for each ERP component of interest between stimulus or task conditions. Analysis of variance (ANOVA) is generally the most common statistical analysis used in ERP research, but non-parametric methods can also be used (e.g., [100]). Following the statistics, it is also recommendable to take good care of the visualization of the results. Journal papers typically report *grand averages* across all participants to show the general characteristics of ERP components obtained for

each experimental condition, while in some cases showing the difference between grand-averaged ERP waveforms among conditions can also be effective (by means of subtraction).

Single-channel analysis has its limitations, however, since the amplitudes of each channel may vary with the setting of the reference (for a review see [101]). *Multi-channel comparisons* of spatial patterns, or the topographic shapes of the potential distribution on the scalp, are more commonly performed [44,102]. One typical approach to spatial pattern analysis is *tomographic analysis*, which estimates the source activity of the observed current [103]. This estimation can facilitate comparisons between EEG task conditions and results from other studies with functional magnetic resonance imaging (fMRI; e.g., [104]).

Over the past decade, a number of alternative EEG analyses have been proposed, such as time-frequency analysis (for a review see [105]), classification analyses (e.g., [106]), and computational model-based analyses [107]. The latter is a hypothesis-oriented approach to the processes in the brain, which has been developed mainly in fMRI research. An advantage of the model-based analysis is that the model can connect brain responses with behavioral responses of the participant. Although there are few model-based ERP studies so far (e.g., [108,109]), the model-based analysis could develop into an important method to understand the relation between the ERP waveform and the cognitive processes that it reflects.

3.3. Final Remarks

Recent trends in ERP research have shown an upsurge in technical developments that allow the combination of ERP measurements with other brain imaging techniques such as fMRI and near-infrared spectroscopy. The classic combination of ERP measurements with behavioral responses, however, is still a very popular one. EEG equipment is relatively cheap and numerous types of auditory stimuli and paradigms have been developed and employed over the years. Some of these were discussed in this overview, but surely more research opportunities are available. Since auditory ERP research enables the researcher to identify neural correlates of sound processing with great temporal acuity, EEG is an ideal tool to analyze our brain's response during the processing of virtually any kind of sound, ranging from simple tone bursts to complex stimuli made up of speech or music.

ACKNOWLEDGEMENTS

Supported by Grant-in-Aid for JSPS Fellows (25-6091) from the Japan Society for the Promotion of Science (to EH), Grant-in-Aid for the 21st Century COE program from the Ministry of Education, Culture, Sports, Science and Technology, Grants-in-Aid for Scientific Research

(14101001, 19103003, 22591276, and 25242002), and the Kyushu University Project for Interdisciplinary Research of Perception and Cognition. We would like to thank Kazuo Ueda for valuable comments on the manuscript.

REFERENCES

- [1] H. Berger, "The electro-encephalogram of man," *Arch. Psychiatr. Nervenkr.*, **108**, 407–431 (1938).
- [2] P. Gloor, "Berger lecture. Is Berger's dream coming true?" *Electroencephalogr. Clin. Neurophysiol.*, **4**, 253–266 (1994).
- [3] P. A. Davis, "The electrical response of the human brain to auditory stimuli," *Am. J. Physiol.*, **126**, 475–476 (1939).
- [4] J. A. Hämäläinen, T. K. Guttorm, U. Richardson, P. Aiku, H. Lyytinen and P. H. Leppänen, "Auditory event-related potentials measured in kindergarten predict later reading problems at school age," *Dev. Neuropsychol.*, **38**, 550–566 (2013).
- [5] E. J. Golob, J. M. Ringman, R. Irimajiri, S. Bright, B. Schaffer, L. D. Medina and A. Starr, "Cortical event-related potentials in preclinical familial Alzheimer disease," *Neurology*, **73**, 1649–1655 (2009).
- [6] I. Dinstein, D. J. Heeger, L. Lorenzi, N. J. Minshew, R. Malach and M. Behrmann, "Unreliable evoked responses in autism," *Neuron*, **75**, 981–991 (2012).
- [7] G. D. Dawson, "A summation technique for the detection of small evoked potentials," *Electroencephalogr. Clin. Neurophysiol.*, **6**, 65–84 (1954).
- [8] R. Galambos and G. C. Sheatz, "An electroencephalograph study of classical conditioning," *Am. J. Psychol.*, **203**, 173–184 (1962).
- [9] S. J. Luck and E. K. Kappenman, Eds., *The Oxford Handbook of Event-Related Potential Components* (Oxford University Press, New York, 2012).
- [10] B. Kotchoubey, "Event-related potentials, cognition, and behavior: A biological approach," *Neurosci. Biobehav. Rev.*, **30**, 42–65 (2006).
- [11] A. M. L. Coenen, "Neuronal activities underlying the electroencephalogram and evoked potentials of sleeping and waking: Implications for information processing," *Neurosci. Biobehav. Rev.*, **19**, 447–463 (1995).
- [12] E. Donchin, W. Ritter and W. C. McCallum, "Cognitive psychophysiology: The endogenous component of the ERP," in *Event Related Brain Potentials in Man*, E. Callaway, P. Teuting and S. H. Koslow, Eds. (Academic Press, New York, 1978).
- [13] A. P. Fonaryova Key, G. O. Dove and M. I. Maguire, "Linking brainwaves to the brain: An ERP primer," *Dev. Neuropsychol.*, **27**, 183–215 (2005).
- [14] J. L. Stone, M. Calderon-Arnulphi, K. S. Watson, K. Patel, N. S. Mander, N. Suss, J. Fino and J. R. Hughes, "Brainstem auditory evoked potentials—A review and modified studies in healthy subjects," *J. Clin. Neurophysiol.*, **26**, 167–175 (2009).
- [15] H. Pratt, A. Starr, H. J. Michalewski, N. Bleich and N. Mittelman, "The auditory P50 component to onset and offset of sound," *Clin. Neurophysiol.*, **119**, 376–387 (2008).
- [16] J. C. Eccles, *The Inhibitory Pathways of the Central Nervous System. The Sherrington Lectures IX* (Liverpool University Press, Liverpool, 1969).
- [17] L. E. Adler, E. Pachtman, R. D. Franks, M. Pecevich, M. C. Waldo and R. Freedman, "Neurophysiological evidence for a defect in neuronal mechanisms involved in sensory gating in schizophrenia," *Biol. Psychiatry*, **17**, 639–654 (1982).
- [18] P. White and C. Yee, "P50 sensitivity to physical and psychological state influences," *Psychophysiology*, **43**, 320–328 (2006).
- [19] M. Hyde, "The N100 response and its applications," *Audiol. Neurootol.*, **2**, 281–307 (1997).
- [20] N. M. Gage and T. P. Roberts, "Temporal integration: Reflections in the M100 of the auditory evoked field," *NeuroReport*, **11**, 2723–2726 (2000).
- [21] S. A. Hillyard, R. F. Hink, V. L. Schwent and T. W. Picton, "Electrical signs of selective attention in the human brain," *Science*, **182**, 177–180 (1973).
- [22] A. Brancucci, "Neural correlates of cognitive ability," *J. Neurosci. Res.*, **90**, 1299–1309 (2012).
- [23] S. A. Hillyard and T. E. Picton, "Electrophysiology of cognition," in *Handbook of Physiology*, F. Plum, Ed. (William and Wilkins, Baltimore, 1987).
- [24] C. D. Yingling and G. E. Nethercut, "Evoked responses to frequency shifted tones: Tonotopic and contextual determinants," *Int. J. Neurosci.*, **22**, 107–118 (1983).
- [25] S. Paulmann, M. Bleichner and S. A. Kotz, "Valence, arousal, and task effects in emotional prosody processing," *Front. Psychol.*, **4**, 345, doi: 10.3389/fpsyg.2013.00345 (2013).
- [26] T. Kikuchi, G. B. Remijn, K. Ueda and Y. Nakajima, "Mismatch negativity related to the perception of whispered speech," *Proc. Aud. Res. Meet. Acoust. Soc. Jpn.*, **43**, 31–35 (2013).
- [27] R. Näätänen, A. W. K. Gaillard and S. Mäntysalo, "Early selective-attention effect on evoked potential reinterpreted," *Acta Psychol.*, **42**, 313–329 (1978).
- [28] T. W. Picton and S. A. Hillyard, "Human auditory evoked potentials. II: Effects of attention," *Electroencephalogr. Clin. Neurophysiol.*, **36**, 191–199 (1974).
- [29] M. Sams, P. Paavilainen, K. Alho and R. Näätänen, "Auditory frequency discrimination and event-related potentials," *Electroencephalogr. Clin. Neurophysiol.*, **62**, 437–448 (1985).
- [30] W. Ritter, P. Paavilainen, J. Lavikainen, K. Reinikainen, K. Alho, M. Sams and R. Näätänen, "Event-related potentials to repetition and change of auditory stimuli," *Electroencephalogr. Clin. Neurophysiol.*, **83**, 306–321 (1992).
- [31] D. H. Loewy, K. B. Campbell and C. Bastien, "The mismatch negativity to frequency deviant stimuli during natural sleep," *Electroencephalogr. Clin. Neurophysiol.*, **98**, 493–501 (1996).
- [32] G. B. Remijn and Y. Sugita, "Mismatch between anticipated and actually presented sound stimuli in humans," *Neurosci. Lett.*, **202**, 169–172 (1996).
- [33] P. R. Hill, G. M. McArthur and D. V. Bishop, "Phonological categorization of vowels: A mismatch negativity study," *NeuroReport*, **15**, 2195–2199 (2004).
- [34] R. Näätänen, P. Paavilainen, T. Rinne and K. Alho, "The mismatch negativity (MMN) in basic research of central auditory processing: A review," *Clin. Neurophysiol.*, **118**, 2544–2590 (2007).
- [35] W. S. Pritchard, S. A. Shappell and M. E. Brandt, "Psychophysiology of N200/N400: A review and classification scheme," *Adv. Psychophysiol.*, **4**, 43–106 (1991).
- [36] C. Alain, S. R. Arnott and T. W. Picton, "Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related potentials," *J. Exp. Psychol. Hum. Percept. Perform.*, **27**, 1072–1089 (2001).
- [37] S. Sutton, M. Braren, J. Zubin and E. R. John, "Evoked potential correlates of stimulus uncertainty," *Science*, **150**, 1187–1188 (1965).
- [38] E. Donchin, D. Karis, T. R. Bashore, M. G. H. Coles and G. Gratton, "Cognitive psychophysiology and human information processing," in *Psychophysiology: Systems, Processes,*

- and Applications, M. G. H. Coles, E. Donchin and S. W. Porges, Eds. (The Guilford Press, New York, 1986).
- [39] M. Kutas and S. A. Hillyard, "Reading senseless sentences: Brain potentials reflect semantic incongruity," *Science*, **207**, 203–205 (1980).
- [40] M. Kutas and K. D. Federmeier, "Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP)," *Annu. Rev. Psychol.*, **62**, 621–647 (2011).
- [41] W. Grey Walter, R. Cooper, V. J. Aldridge, W. C. McCallum and A. L. Winter, "Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain," *Nature*, **203**, 380–384 (1964).
- [42] T. Mitsudo, C. Gagnon, H. Takeichi and S. Grondin, "An electroencephalographic investigation of the filled-duration illusion," *Front. Integr. Neurosci.*, **5**, 84, doi: 10.3389/fnint.2011.00084 (2012).
- [43] S. Chennu, V. Noreika, D. Gueorguiev, A. Blenkman, S. Kochen, A. Ibáñez, A. M. Owen and T. A. Bekinschtein, "Expectation and attention in hierarchical auditory prediction," *J. Neurosci.*, **33**, 11194–11205 (2013).
- [44] T. W. Picton, S. Bentin, P. Berg, E. Donchin, S. A. Hillyard, R. Johnson Jr., G. A. Miller, W. Ritter, D. S. Ruchkin, M. D. Rugg and M. J. Taylor, "Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria," *Psychophysiology*, **37**, 127–152 (2000).
- [45] S. J. Luck, *An Introduction to the Event-related Potential Technique* (MIT Press, Cambridge, Mass., 2005).
- [46] P. Tallgren, S. Vanhatalo, K. Kaila and J. Voipio, "Evaluation of commercially available electrodes and gels for recording of slow EEG potentials," *Clin. Neurophysiol.*, **116**, 799–806 (2005).
- [47] F. Rösler, M. Heil and E. Hennighausen, "Distinct cortical activation patterns during long-term memory retrieval of verbal, spatial, and color information," *J. Cogn. Neurosci.*, **7**, 51–65 (1995).
- [48] D. M. Tucker, "Spatial sampling of head electrical fields: The geodesic sensor net," *Electroencephalogr. Clin. Neurophysiol.*, **87**, 154–163 (1993).
- [49] H. Davis and S. Zerlin, "Acoustic relations of the human vertex potential," *J. Acoust. Soc. Am.*, **39**, 109–116 (1966).
- [50] R. Galambos, S. Makeig and P. J. Talmachoff, "A 40-Hz auditory potential recorded from the human scalp," *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 2643–2647 (1981).
- [51] K. Itoh, Y. Okumiya-Kanke, Y. Nakayama, I. L. Kwee and T. Nakada, "Effects of musical training on the early auditory cortical representation of pitch transitions as indexed by change-N1," *Eur. J. Neurosci.*, **36**, 3580–3592 (2012).
- [52] S. Koelsch, T. Gunter, A. D. Friederici and E. Schröger, "Brain indices of music processing: "Nonmusicians" are musical," *J. Cogn. Neurosci.*, **12**, 520–541 (2000).
- [53] M. Kutas, "Views on how the electrical activity that the brain generates reflects the functions of different language structures," *Psychophysiology*, **34**, 383–398 (1997).
- [54] E. Gontier, E. Hasuo, T. Mitsudo and S. Grondin, "EEG investigations of duration discrimination: The intermodal effect is induced by an attentional bias," *PLoS ONE*, **8**, e74073, doi: 10.1371/journal.pone.0074073 (2013).
- [55] D. Deutsch, "An auditory illusion," *Nature*, **251**, 307–309 (1974).
- [56] J. Ross, M. Tervaniemi and R. Näätänen, "Neural mechanisms of the octave illusion: Electrophysiological evidence for central origin," *NeuroReport*, **8**, 303–306 (1996).
- [57] R. Miyauchi and Y. Nakajima, "The category of 1:1 ratio caused by assimilation of two neighboring empty time intervals," *Hum. Movement Sci.*, **26**, 717–727 (2007).
- [58] T. Mitsudo, Y. Nakajima, G. B. Remijn, H. Takeichi, Y. Goto and S. Tobimatsu, "Electrophysiological evidence of auditory temporal perception related to assimilation between two neighboring time intervals," *NeuroQuantology*, **7**, 114–127 (2009).
- [59] J. C. R. Licklider, J. C. Webster and J. M. Hedlund, "On the frequency limits of binaural beats," *J. Acoust. Soc. Am.*, **22**, 468–473 (1950).
- [60] H. Pratt, A. Starr, H. J. Michalewski, A. Dimitrijevic, N. Bleich and N. Mittelman, "Cortical evoked potentials to an auditory illusion: Binaural beats," *Clin. Neurophysiol.*, **120**, 1514–1524 (2009).
- [61] T. Kuroda, R. Miyauchi, E. Hasuo and Y. Nakajima, "What one should care when conducting auditory psychological experiments," *J. Des.*, **9**, 63–66 (2008) (in Japanese).
- [62] H. Riquimaroux, "Pitfalls for auditory experiments," *J. Acoust. Soc. Jpn. (J)*, **60**, 614–619 (2004).
- [63] W. D. Keidel and M. Spreng, "Neurophysiological evidence for the Stevens power function in man," *J. Acoust. Soc. Am.*, **38**, 191–195 (1965).
- [64] H. Davis, T. Mast, N. Yoshie and S. Zerlin, "The slow response of the human cortex to auditory stimuli: Recovery process," *Electroencephalogr. Clin. Neurophysiol.*, **21**, 105–113 (1966).
- [65] R. Hari, K. Kaila, T. Katila, T. Tuomisto and T. Varpula, "Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: Implications for their neural generation," *Electroencephalogr. Clin. Neurophysiol.*, **54**, 561–569 (1982).
- [66] K. Inui, T. Urakawa, K. Yamashiro, N. Otsuru, M. Nishihara, Y. Takeshima, S. Keceli and R. Kakigi, "Non-linear laws of echoic memory and auditory change detection in humans," *BMC Neurosci.*, **11**, 80 (2010).
- [67] H. H. Jasper, "The ten-twenty electrode system of the International Federation," *Electroencephalogr. Clin. Neurophysiol.*, **10**, 370–375 (1958).
- [68] American Electroencephalographic Society, "American Electroencephalographic Society. Guideline thirteen: Guidelines for standard electrode position nomenclature," *J. Clin. Neurophysiol.*, **11**, 111–113 (1994).
- [69] G. E. Chatrian, E. Lettich and P. L. Nelson, "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity," *Am. J. EEG Technol.*, **25**, 83–92 (1985).
- [70] M. R. Nuwer, G. Comi, R. Emerson, A. Fuglsang-Frederiksen, J. M. Guérit, H. Hinrichs, A. Ikeda, F. J. C. Luccas and P. Rappelsburger, "IFCN standards for digital recording of clinical EEG," *Electroencephalogr. Clin. Neurophysiol.*, **106**, 259–261 (1998).
- [71] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution EEG and ERP measurements," *Clin. Neurophysiol.*, **112**, 713–719 (2001).
- [72] R. J. Croft and R. J. Barry, "Removal of ocular artifact from the EEG: A review," *Clin. Neurophysiol.*, **30**, 5–19 (2000).
- [73] T. W. Picton, P. van Roon, M. L. Armilio, P. Berg, N. Ille and M. Scherg, "The correction of ocular artifacts: A topographic perspective," *Clin. Neurophysiol.*, **111**, 53–65 (2000).
- [74] N. Gosselin, A. Mathieu, S. Mazza, D. Petit, J. Malo and J. Montplaisir, "Attentional deficits in patients with obstructive sleep apnea syndrome: An event-related potential study," *Clin. Neurophysiol.*, **117**, 2228–2235 (2006).
- [75] I. C. Fiebelkorn, J. J. Foxe and S. Molholm, "Dual mechanisms for the cross-sensory spread of attention: How

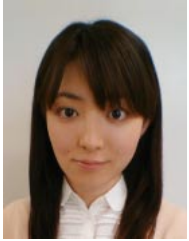
- much do learned associations matter?" *Cereb. Cortex*, **20**, 109–120 (2010).
- [76] T. Aihara, Y. Takeda, K. Takeda, W. Yasuda, T. Sato, Y. Otaka, T. Hanakawa, M. Honda, M. Liu, M. Kawato, M. Sato and R. Osu, "Cortical current source estimation from electroencephalography in combination with near-infrared spectroscopy as a hierarchical prior," *Neuroimage*, **59**, 4006–4021 (2012).
- [77] P. Berg and M. A. Scherg, "Multiple source approach to the correction of eye artifacts," *Electroencephalogr. Clin. Neurophysiol.*, **90**, 229–241 (1994).
- [78] J. S. Snyder, C. Alain and T. W. Picton, "Effects of attention on neuroelectric correlates of auditory stream segregation," *J. Cogn. Neurosci.*, **18**, 1–13 (2006).
- [79] B. Mohr, T. Endrass, O. Hauk and F. Pulvermüller, "ERP correlates of the bilateral redundancy gain for words," *Neuropsychologia*, **45**, 2114–2124 (2007).
- [80] B. R. Zendel and C. Alain, "Enhanced attention-dependent activity in the auditory cortex of older musicians," *Neurobiol. Aging*, **35**, 55–63 (2014).
- [81] S. Onishi and H. Davis, "Effects of duration and rise time of tone bursts on evoked V potentials," *J. Acoust. Soc. Am.*, **44**, 582–591 (1968).
- [82] J. A. Altman and S. F. Vaitulevich, "Auditory image movement in evoked potentials," *Electroencephalogr. Clin. Neurophysiol.*, **75**, 323–333 (1990).
- [83] S. Swink and A. Stuart, "The effect of gender on the N1–P2 auditory complex while listening and speaking with altered auditory feedback," *Brain Lang.*, **122**, 25–33 (2012).
- [84] K. A. Barrett and J. M. Fulfs, "Effect of gender on the mismatch negativity auditory evoked potential," *J. Am. Acad. Audiol.*, **9**, 444–451 (1998).
- [85] K. Kasai, K. Nakagome, A. Iwanami, M. Fukuda, K. Itoh, I. Koshida and N. Kato, "No effect of gender on tonal and phonetic mismatch negativity in normal adults assessed by a high-resolution EEG recording," *Brain Res. Cogn. Brain Res.*, **13**, 305–312 (2002).
- [86] E. Nagy, G. F. Potts and K. A. Loveland, "Sex-related ERP differences in deviance detection," *Int. J. Psychophysiol.*, **48**, 285–292 (2003).
- [87] J. L. Wunderlich, B. K. Cone-Wesson and R. Shepherd, "Maturation of the cortical auditory evoked potential in infants and young children," *Hear. Res.*, **212**, 185–202 (2006).
- [88] A. Pfefferbaum, J. M. Ford, W. T. Roth and B. S. Kopell, "Age-related changes in auditory event-related potentials," *Electroencephalogr. Clin. Neurophysiol.*, **49**, 266–276 (1980).
- [89] K. L. Tremblay, M. Piskosz and P. Souza, "Aging alters the neural representation of speech cues," *NeuroReport*, **13**, 1865–1870 (2002).
- [90] K. L. Tremblay, C. Billings and N. Rohila, "Speech evoked cortical potentials: Effects of age and stimulus presentation rate," *J. Am. Acad. Audiol.*, **15**, 226–237 (2004).
- [91] K. Alho, K. Sainio, N. Sajaniemi, K. Reinikainen and R. Näätänen, "Event-related brain potential of human newborns to pitch change of an acoustic stimulus," *Electroencephalogr. Clin. Neurophysiol.*, **77**, 151–155 (1990).
- [92] N. Kraus, T. McGee, T. Carell, A. Sharma, A. Micco and T. Nicol, "Speech-evoked cortical potentials in children," *J. Am. Acad. Audiol.*, **4**, 238–248 (1993).
- [93] H. Gaeta, D. Friedman, W. Ritter and J. Chen, "An event-related potential study of age-related changes in sensitivity to stimulus deviance," *Neurobiol. Aging*, **19**, 447–459 (1998).
- [94] G. Josse and N. Tzourio-Mazoyer, "Hemispheric specialization for language," *Brain Res. Rev.*, **44**, 1–12 (2004).
- [95] D. Deutsch, "Handedness and memory for tonal pitch," in *Neuropsychology of Left-handedness*, J. Heron, Ed. (Academic Press, New York, 1980).
- [96] R. C. Oldfield, "The assessment and analysis of handedness: The Edinburgh inventory," *Neuropsychologia*, **9**, 97–113 (1971).
- [97] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, **13**, 411–430 (2000).
- [98] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, **37**, 163–178 (2000).
- [99] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J. Neurosci. Methods*, **158**, 300–312 (2006).
- [100] E. Maris and R. Oostenveld, "Nonparametric statistical testing of EEG- and MEG-data," *J. Neurosci. Methods*, **164**, 177–190 (2007).
- [101] M. M. Murray, D. Brunet and C. M. Michel, "Topographic ERP analyses: A step-by-step tutorial review," *Brain Topogr.*, **20**, 249–264 (2008).
- [102] G. McCarthy and C. C. Wood, "Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models," *Electroencephalogr. Clin. Neurophysiol.*, **62**, 203–208 (1985).
- [103] C. M. Michel, M. M. Murray, G. Lantz, S. Gonzalez, L. Spinelli and R. Grave de Peralta, "EEG source imaging," *Clin. Neurophysiol.*, **115**, 2195–2222 (2004).
- [104] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details," *Methods Find. Exp. Clin. Pharmacol.*, **24**, 5–12 (2002).
- [105] S. Makeig, S. Debener, J. Onton and A. Delorme, "Mining event-related brain dynamics," *Trends Cogn. Sci.*, **8**, 204–210 (2004).
- [106] J. N. Mak, Y. Arbel, J. W. Minett, L. M. McCane, B. Yuksel, D. Ryan, D. Thompson, L. Bianchi and D. Erdogmus, "Optimizing the P300-based brain-computer interface: Current status, limitations and future directions," *J. Neural Eng.*, **8**, 025003, doi: 10.1088/1741-2560/8/2/025003 (2011).
- [107] J. P. O'Doherty, A. Hampton and H. Kim, "Model-based fMRI and its application to reward learning and decision making," *Ann. N.Y. Acad. Sci.*, **1104**, 35–53 (2007).
- [108] R. B. Mars, S. Debener, T. E. Gladwin, L. M. Harrison, P. Haggard, J. C. Rothwell and S. Bestmann, "Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise," *J. Neurosci.*, **28**, 12539–12545 (2008).
- [109] F. Karayanidis, S. Jamadar, H. Ruge, N. Phillips, A. Heathcote and B. U. Forstmann, "Advance preparation in task-switching: converging evidence from behavioral, brain activation, and model-based approaches," *Front. Psychol.*, **1**, 25, doi: 10.3389/fpsyg.2010.00025 (2010).



Gerard Bastiaan Remijn (b. 1970) received a BA and MA in Psychology from Leiden University, the Netherlands, and a Ph.D. from the Kyushu Institute of Design. His research mainly focuses on cognitive psychology and neuroscience. He is currently an Associate Professor at the Human Science Department/Department of Acoustic Design at Kyushu University.



Emi Hasuo (b. 1983) received a Bachelor's degree from Kunitachi College of Music, and a Master's and a Ph.D. degree from Kyushu University. She went on to post-doctoral training at Laval University and is currently a research fellow of JSPS. Her research has so far focused on understanding the basic mechanisms underlying human time and rhythm perception.



Haruna Fujihira (b. 1981) received a Bachelor of Health Science degree from Okayama University, Japan, in 2003, and the Bachelor of Design and Master of Design degrees from Kyushu University, Japan, in 2011 and 2013. She is currently a doctoral student at Kyushu University. Her research interests include auditory evoked potentials in elderly adults.

Satoshi Morimoto (b. 1982) holds a Masters degree in Science. From 2011–2013, he worked at the Advanced Telecommunications Research Institute International (ATR) as a technical researcher. He is currently studying at the Graduate School of Design, Kyushu University. His interests are computational mechanisms underlying perceptual and cognitive functions.