Aus dem Institut für Medizinische Psychologie
der Universität Tübingen

The influence of voice familiarity on the Mismatch Response in two- to four-month-old infants

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Thöne, Leonie

2019

Dekan: Professor Dr. I. B. Autenrieth

1. Berichterstatter: Professor Dr. J. Born

2. Berichterstatter: Professor Dr. M Wilke

Tag der Disputation: 25.01.2019

Contents

List of Figures

List of Tables

Abbreviations

ANOVA: Analysis of Variance
CTR: Conventional Trial Rejection
EEG: Electroencephalography
EKP: Ereignis-Korreliertes Potential
EN: Early Negativity
EOG: Electrooculography
ERP: Event Related Potential
FA: Flächenamplitude
fMRI: Functional Magnetic Resonance Imaging
FN: Frühe Negativität
FTPV: Fronto-Temporal Positivity to Voices
ICR: Independent Channel Rejection
ISI: Interstimulus Interval
LN: Late Negativity
M: Mean
MMN: Mismatch Negativity
MMR: Mismatch Response
N: Negative
P: Positive
PC: Positive Component
PK: Positive Komponente
PSP: Postsynaptic Potential
ROI: Region of Interest
SEM: Standard Error of the Mean
SD: Standard Deviation
VSR: Voice-Specific Response

1 INTRODUCTION

This introduction sheds light on the preattentive and cognitive processes that reflect voice recognition as well as on electrophysiological markers of voice processing with an emphasis on their development during ontogeny. For a better understanding of the dynamics that are related to voice recognition throughout development, an overview of the current scientific view of voice recognition and processing in adults is given. Subsequently, a comparison to studies on auditory recognition memory conducted with developmental population is drawn, focusing on the timespan from birth to the first months of life.

Furthermore, the electrophysiological markers that we used in the present study are outlined. An introduction to auditory event-related potentials (ERPs) and to the component that was targeted in our study follows: the mismatch negativity (MMN) as it is called in adults or the mismatch response (MMR), the term that is more frequently used for it infant correlate. Finally, some electrophysiological studies that explored neural correlates of voice familiarity are described. In previous infant ERP studies, characteristics of adult ERP components are often used to interpret infant components. For a better understanding of these comparisons, an insight into the ‘adult counterparts’ of the targeted ERP components is given.

1.1 Voice recognition in adults

Voice recognition is a process that enables the listener to gain paralinguistic information on the speaker, irrespectively of the speech content that is being transmitted. Thus, recognizing familiar voices is considered as an important aspect of social interaction in human species. Voice processing happens in different cognitive categories: through characteristic voice patterns, we can differentiate between a familiar or an unfamiliar voice, evaluate affective states and determine the speaker’s gender independently from visual information. Voice processing furthermore enables identification of the speaker through constant voice characteristics that are transmitted by their vocalizations (Belin, Fecteau, & Bedard, 2004).

Voice recognition is a part of auditory recognition memory that itself is a type of declarative memory. Information on acoustic perception is processed in the primary auditory cortex that is networked to the higher-order auditory association cortices. From these points, projections to the hippocampus and the rhinal cortex through the parahippocampal cortex have been found. Neural substrates that have been proposed to be part of the formation of auditory recognition memory are illustrated in Figure 1.1 (deRegnier, 2007; Squire, Schmolck, & Stark, 2001).

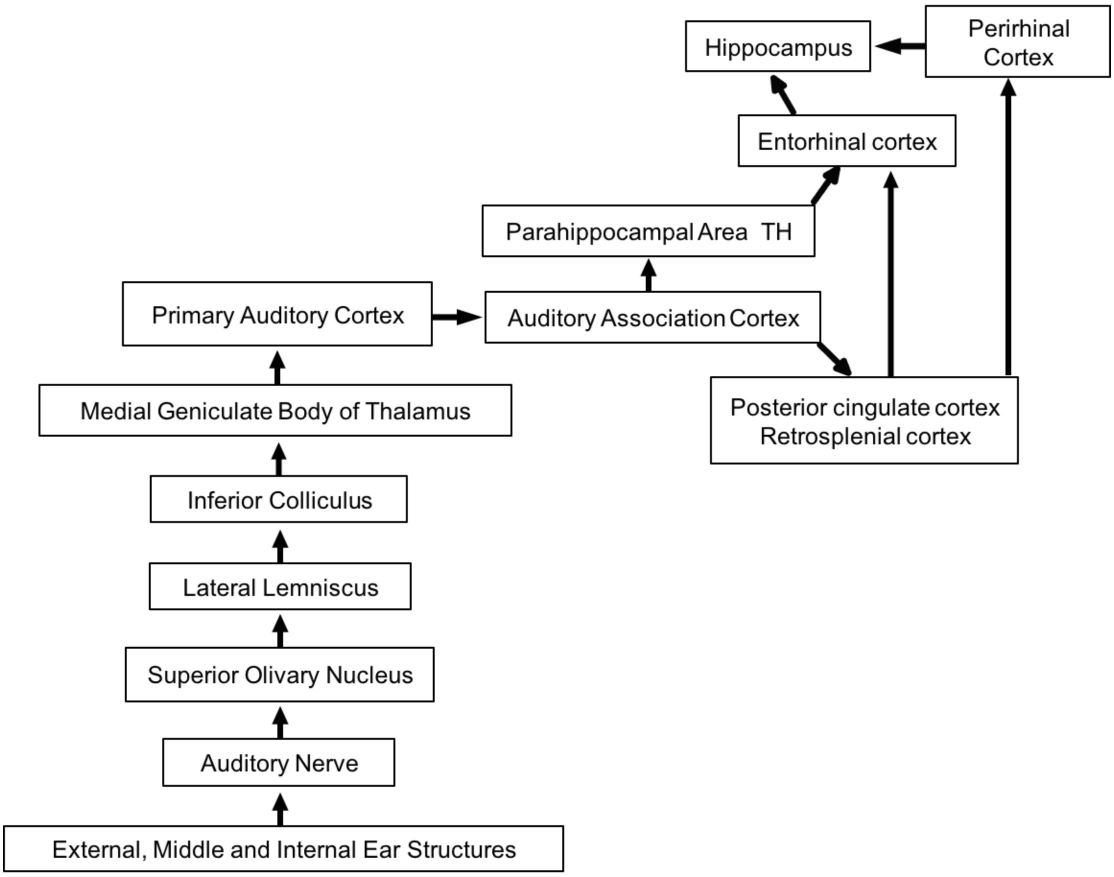

Figure 1.1: Neural substrates of auditory recognition memory.

Neural structures that are assumed to be involved in acoustic perception as well as auditory recognition memory and their connections between each other. Figure adapted from *Auditory recognition memory in infancy* (p.147) by R. deRegnier, in *Infant EEG and event-related potentials* edited by M. de Haan, 2007, Psychology Press. Adapted and printed with permission of Taylor and Francis Group.

The networks that process voices differently from other non-vocal sounds within the human auditory recognition system have been a subject matter of recent research. There is a growing body of evidence that the human brain exhibits specific neural pathways for voice processing within the auditory system. Voice-selective regions within the auditory cortex which are located bilaterally in mid and anterior parts of the superior temporal gyrus (‘temporal voice areas’) seem to play a crucial role in voice processing (Belin, Bestelmeyer, Latinus, & Watson, 2011). In a functional magnetic resonance image (fMRI) study stimuli of vocal origin provoked an increased activation of these areas in comparison to sounds of non-vocal origin. These results correspond with the existence of cortical regions that respond selectively to voices (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000).

Levy and colleagues described a possible electrophysiological correlate to the assumed voice-specific pathways. They used ERPs to evaluate responses elicited by non-vocal (instruments playing) and vocal stimuli (singer singing). A positive peaking at 320 ms after stimulus onset was observed in the ERPs evoked exclusively by human voices that they described as ‘voice-specific response’ (VSR) (Levy, Granot, & Bentin, 2001).

Besides the VSR, a voice-specific frontotemporal positivity (FTPV) peaking at about 200 ms latency has been found in adults (Charest et al., 2009) and also in children between the age of four and five years predominating at the right temporal sites (Rogier, Roux, Belin, Bonnet-Brilhault, & Bruneau, 2010). In a fMRI study, a corresponding response was found and located in the vicinity of the aforementioned temporal voice areas (Capilla, Belin, & Gross, 2013). It was proposed that the early FTPV reflects the ‘what’ pathway in auditory processing and a comparison to the face-preferential N170 in visual ERPs was drawn (Charest et al., 2009).

The ability to identify a speaker by their voice does not seem to be confined to the human species. The research team of Scott and colleagues discovered that macaques also show higher activity in the fMRI in voice-sensitive brain areas (anterior superior temporal plane) while listening to vocalizations of familiar conspecifics. These findings suggest that recognizing the speaker’s or conspecific identity is an evolutionarily conserved brain function (Scott, 2008).

Voices are often considered as an ‘auditory face’ and similarities between the neural processing of face and voice recognition have been drawn. Belin and colleagues introduced a model for voice processing based on analogs to the model of face perception postulated by Bruce and Young (Belin et al., 2011; Bruce & Young, 1986). According to the former, voice processing runs through a low-level analysis that is realized in sub-cortical nuclei and core regions of the auditory cortex. Subsequently, a voice-specific structural analysis occurs that encompasses the extraction and division into three main types of voice information that are further processed in partly interacting functional pathways. They postulated that there exist 1) a pathway for vocal speech analysis, 2) a pathway for vocal affect analysis and 3) a pathway to determine vocal identification which is constructed of voice recognition units, each of them corresponding to a different familiar voice. The latter is assumed to be localized in the region of the right anterior superior temporal sulcus (Belin et al., 2011).

The process of voice identification is subdivided into two cognitive skills that are related to different cognitive mechanisms. ‘Voice recognition’ describes the ability to recognize familiar voices, whereas ‘voice discrimination’ is defined as the mechanism to differentiate voices from each other. Lesion studies underpinned the hypothesis of two independent cognitive processes. Patients with lesions located in either temporal lobes showed impaired voice discrimination, while lesions in the right temporal lobe involved the inability to distinguish whether a voice is familiar or unfamiliar to the listener (Van Lancker, Cummings, Kreiman, & Dobkin, 1988; Van Lancker & Kreiman, 1987), defined as the term ‘phonagnosia’ by Van Lancker & Canter (1982).

1.2 Ontogeny of auditory recognition memory

As mentioned above, voice familiarity is a part of the auditory recognition memory. There is a growing body of evidence that recognition memory has its origin in early ontogeny and reflects one of the first cognitive functions (deRegnier, 2007). Sensory recognition memory, especially in the auditory modality, is considered to be one of the earliest types of explicit memory (Jabès & Nelson, 2015). It is poorly understood how the aforementioned neuroanatomic substrates of auditory recognition memory develop throughout life. However, it was found that the development of the human hippocampus initiates *in utero* and the majority of its neuronal formation is completed in the first half of gestation (Seress, Abraham, Tornoczky, & Kosztolanyi, 2001). These findings support the idea that the tools for recognition memory develop even before birth (deRegnier, 2007).

Fetal behavioral studies are in line with the idea that recognition memory has its origins in early ontogeny. An example of this assumption provided Kisilevsky and colleagues who found that fetuses habituate to a repeated sound and dishabituate to its novel presentation at the end of gestation. The habituation response was quantified by the fetal heart rate (Kisilvesky & Muir, 1991).

Sound experience in humans, including voice experience, begins already in prenatal life. Sounds are conducted through the tissues of the maternal abdominal wall, the uterus as well as the amniotic fluid and perceived by the fetus (Querleu, Renard, Versyp, Paris-Delrue, & Vervoort, 1988). The soundscape that the fetus is surrounded by includes external noises and sounds produced by the mother and her body: voices, heartbeats, digestive sounds and blood flow (deRegnier, 2007). Auditory sense develops during the fetal period from the 28th gestational week on and at the time of birth the neonate commands moths of acoustic experience acquired during their stay in utero (Querleu, Renard, Versyp, Paris-Delrue, & Crepin, 1988).

Since the uterus serves as a kind of filter for airborne sounds, mainly low frequencies are perceived *in utero*. Due to the conductance through the maternal tissues, voice characteristics are audible as a muffled version. However, the acoustic properties of the mother's voice do not change as much as other voices, since transmission of her voice additionally occurs through her body tissues. Moreover, the maternal voice is perceived louder than other outside voices by the fetus (deRegnier, 2007; Lecanuet & Schaal, 1996). Studies showed that the mother's voice has an impact on the fetal heart rate, either decreasing (Fifer & Moon, 1995) or accelerating it in comparison to a stranger's voice (Kisilevsky et al., 2003). Fifer and colleagues suggested that the maternal voice influences structuring processes within the maturing brain (Fifer & Moon, 1994).

To sum up, the formation of recognition memory seems to depend on both the state of maturity and the amount of prior sound experience during the perinatal period. However, deRegnier suggested that intensive exposure to a specific stimulus is required (as it occurs naturally with the maternal voice *in utero*) to form stable long-term memory templates in the neonatal period (deRegnier, 2007).

1.2.1 Role of the mother's voice in infancy

Familiarity with the maternal voice has been a topic of several behavioral studies in the past. Being perceived by the fetus before birth as described before, the mother's voice continues to be a highly familiar and constant acoustic stimulus throughout development. Preference for the maternal voice persists after birth in the human species. In the first hours of their lives, newborns show reactions (suckling behavior, orienting response) superior to their mother's voice than to an unfamiliar one (DeCasper & Fifer, 1980; Querleu et al., 1984). Fifer and Moon found that newborns prefer a filtered version of the maternal voice that imitates the *in utero* conditions over the unfiltered *ex utero* version (Fifer & Moon, 1995). In contrast, newborns seem to have the ability to discriminate the father's voice between unfamiliar male voices, but do not show a preference for it (DeCasper & Prescott, 1984).

Recognizing the maternal voice among others is suggested to play a substantial role in early social communication abilities (Abrams et al., 2016). Moreover, familiarity with the mother's voice might provide evolutionary advantages as a behavioral study in other mammals showed. Charrier and colleagues investigated how fur seal pups identify their mothers by their vocalizations after weeks of separation due to foraging trips that the mothers undertake to feed their offspring. It was shown that seal pups react to their mother's characteristic vocalizations two to five days after birth and can recognize them amongst other seal females after two to three weeks of separation (Charrier, Mathevon, & Jouventin, 2001).

Deducing from the observations made on the familiarity with the maternal voice in fetal life and infancy, it does not only seem to play a crucial role in the mother-infant interaction but also in early learning processes during ontogeny. Due to the assumption that the mother's voice represents an important cue for the formation of auditory recognition memory in early infancy, we chose it to play a key part in the present study.

Other researchers had used the mother's as a highly familiar acoustic stimulus in comparison to a female stranger's voice in previous ERP studies that investigated the formation of auditory recognition memory (Beauchemin et al., 2011; deRegnier, 2007; deRegnier, Long, Georgieff, & Nelson, 2007; Mai et al., 2012; Purhonen, Kilpelainen-Lees, Valkonen-Korhonen, Karhu, & Lehtonen, 2005).

1.2.2 Voice processing in early childhood

Few electrophysiological and imaging studies on the development of voice processing in infancy have been realized. Grossmann and collogues conducted a near-infrared spectroscopy (NIRS) study in which they compared oxygenated hemoglobin concentration in distinct brain regions of the infant brain as a response to either vocal (speech and non-speech vocalizations) and non-vocal stimuli. Greater hemodynamic responses to vocal stimuli than to non-vocal stimuli were detected bilaterally in the superior temporal cortex in seven-month-old infants. However, a similar voice-sensitive response could not be detected in four-month-old infants. These findings suggest that the temporal voice-sensitive areas, as they have been described in adults, emerge between the age of four and seven months (Grossmann, Oberecker, Koch, & Friederici, 2010).

The absence of voice-specific regions in early infancy as it was shown by Grossmann and colleagues seems to object the ability of the neonate to recognize their own mother's voice among others as it had been found in several behavioral studies. However, cues for immature networks involved in voice processing in early childhood are provided by studies utilizing electrophysiological methods, such as ERPs. It is likely that mechanisms of the voice recognition memory are different from the processes that lie beyond voice processing in adults and older children (Mai et al., 2012). In fact, Beauchemin postulated that infants process their mothers' voices at first preattentively and in the course of development process it at a cognitive level (Beauchemin et al., 2011).

The auditory MMN (Näätänen & Michie, 1979) as an ERP component has been used to assess auditory familiarity. Its infant counterpart, the MMR, is suggested to be a promising tool for measuring voice familiarity in early infancy, since it reflects preattentive processes (Beauchemin et al., 2011; Cheour, 2007). As the MMR is elicited independently of the listener's attention, it further has proved to be an appropriate electrophysiological marker in developmental populations (Cheour, 2007; Näätänen, 2000).

To gain a better understanding of the neural processes behind the MMN and its infant analog, the ERP component we targeted in the study, a short introduction on ERPs as well as an insight into different auditory ERP components follow.

1.3 Event-related potentials

Even-related potentials (ERPs) are scalp-derived potentials that are elicited by a time-locked event, such as a sensory stimulus. Since the measured electroencephalogram (EEG) signal usually encompasses a mixture of EEG oscillations, noises and other components, ERPs only become visible when averaging many single potentials (epochs) in response to a repetitive exposure of the targeted stimulus. Neurophysiologic investigations showed that ERPs reflect postsynaptic potentials (PSPs) forming a directed dipole. When PSPs are evoked in a group of similarly orientated neurons simultaneously, their dipoles sum together. The summed dipole is conducted on the scalp surface and can be derived as specific potentials (for a review see Luck, 2014b, Chap. 2).

ERP components are traditionally categorized into three different types: (1) exogenous sensory components are elicited obligatorily when a subject is exposed to a stimulus, (2) endogenous components are elicited task-specifically and thus are not obligatory in the response to a stimulus (3) motor components that reflect preparation and execution of a motor response (for review see Luck, 2014b, Chap. 3).

1.4 Auditory ERP components

1.4.1 Adults

Auditory ERPs are electrophysiological responses to an acoustic stimulation reflecting neural processes that occur when we are confronted with an acoustic stimulus. The full length of neural pathway activation in response to a sound can be retraced in different auditory ERP components (for review on auditory ERPs see Pratt, 2012; Trainor, 2007). Each component is related to the activation and synchronization of a specific group of neurons within in the auditory system, beginning with the receptor potential within the cochlear, along the auditory nerve, the brainstem, through the thalamus to the auditory cortices.

By means of their latency, auditory ERP responses are divided into (1) auditory nerve and brainstem response measured at 10 ms after stimulus onset; (2) the auditory middle-latency response detected between 40-50 ms; and (3) the long-latency response defined as components occurring from 50 ms on post-stimulus (see Figure 1.2.).

Long-latency auditory potentials are detected with a bandpass of 0.1-100 Hz. They encompass several negative and positive components initiating with a positive peak at around 50ms (P50 or P1) followed by a negative peak at 100 ms (N100 or N1). N1 is followed by a second positive peak, often referred to as P2 occurring with a latency around 180 ms in adults. P1, N1, and P2 are categorized as exogenous components since they do not require a task to be observed. The P2 component is followed by several positive and negative peaks (N200, P300, etc.) that are presumably not related to processes on the sensory perception level but are elicited when the presented sound is connected to a specific task.

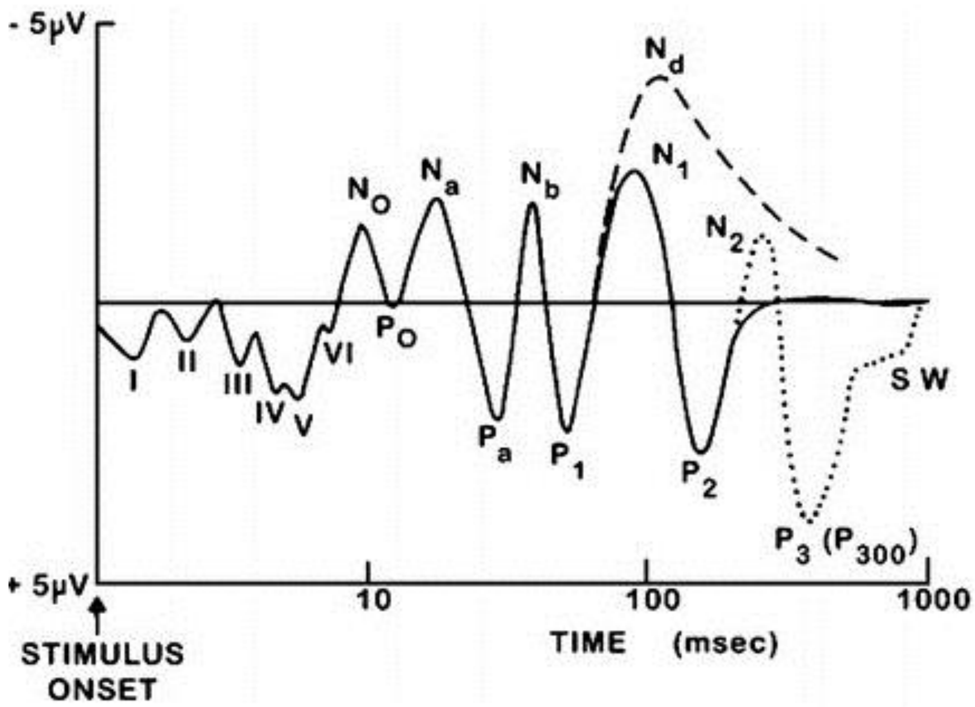

Figure 1.2: Auditory event-related potentials.

Schematic depiction of auditory event-related potentials and their latencies. The first five peaks reflect early brainstem potentials (I-VI), followed by the middle-latency potentials (N0-Nb) and the long-latency potentials (N1-P2). Furthermore, non-obligatory task- and attention-related components are displayed: Nd (dashed line), N2, P300 and slow wave (SW) (pointed line). Figure from *Electrophysiology of cognitive processing* (p.35) by Hillyard S.A. and Kutas M. in *Annual Review of Psychology* (p.33-61), 1983.

1.4.2 Early infancy

The long-latency auditory ERPs and its components undergo different changes during ontogeny. However, the dynamics of the changes in ERP components throughout maturation remain mostly in the dark. In general, latencies, the span between stimulus onset and the detected ERP component, tend to shorten (Jing & Benasich, 2006) and amplitudes seem to increase throughout childhood (Cheour, 2007). The changes that occur in the morphology of ERP components in a relatively short time span during ontogeny might lead to a high inter-individual variability frequently found in developing test populations (Coch & Gullick, 2012; deRegnier, 2007). One theory says that these developmental changes in ERP characteristics are a result of increasing neural signal conduction due to advancing myelination (Eggermont, 1992) and, in the case of auditory components, a growing number of synapses in the auditory cortex (deRegnier, 2007;

Huttenlocher, De Courten, Garey, & Van der Loos, 1982). In fact, in a neurohistological study, synaptogenesis was observed to reach its peak in the auditory cortex at the age of three months. Synapses were found to be eliminated in the auditory cortex until the age of twelve years (Huttenlocher & Dabholkar, 1997).

The complex of alternating positive and negative waves as mentioned above has only been observed in adults and auditory ERP components seem to differentiate throughout brain maturation. In fact, the processing of acoustic information is presumably slower in infants than in adults due to poorer axonal myelination and lower cortical linkage (Dehaene-Lambertz et al., 2010).

If an equivalent of the N1 component exists in early infancy, has been topic in infant ERP research. For example, it has been discussed if a broader Interstimulus Interval (ISI) favors elicitation of N1 in infancy (Wunderlich, Cone-Wesson, & Shepherd, 2006). Wunderlich and colleagues found that an N1 is reliably evoked from birth on and increases with aging (2006). Other studies suggest that N1b, a subcomponent, solely is constantly elicited from school age or adolescence on (Albrecht, Suchodoletz, & Uwer, 2000; Ponton, Eggermont, Kwong, & Don, 2000).

In the neonatal period, a broad positive wave dominates the response to tones or speech sounds, peaking frontocentrally at 250-300 ms that is often compared to the adult P2 component. (Cheour, 2007; deRegnier, Wewerka, Georgieff, Mattia, & Nelson, 2002; Kurtzberg, Hilpert, Kreuzer, & Vaughan, 1984; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002; Trainor, 2007). Between two and four months, the infant P2 component still prevails among other exogenous auditory ERP components (Mai et al., 2012; Purhonen et al., 2005; Trainor et al., 2003).

1.5 Auditory deviance detection response

1.5.1 Mismatch negativity (MMN) in adults

Some non-obligatory ERP components have been discussed to reflect auditory deviance detection in the human brain. One of them, the auditory MMN, is often described as a change-specific response to a deviant acoustic stimulation in a stream of consecutive frequent stimuli (for a review on MMN see Näätänen, 2012; Näätänen & Picton, 1987; Winkler, Karmos, & Näätänen, 1996).

A typical MMN experiment involves an oddball sequence that is played to the test person. Characteristically, this oddball paradigm consists of frequent repetitive stimuli (‘standard stimuli’) disrupted by infrequent rare stimuli (‘deviant stimuli’) differing in sound features from the standard stimuli. However, certain prerequisites have to be fulfilled: at least three standard stimuli have to precede the deviant stimulus to elicit an MMN response (Cheour, Leppanen, & Kraus, 2000).

The standard stimulus within an oddball sequence of acoustic stimuli typically evokes late-latency auditory ERP components, one peaking at around 100 ms (N1) and another positive one peaking at 180-200 ms (P2) after stimulus onset. An additional negative response, the MMN, between 100 and 250 ms is elicited by the deviant stimulus beginning at the peak of N1 and overlapping P2. For better visualization, the MMN response is displayed as a difference wave that is computed by subtracting the original ERP response to the frequent stimulus from the original ERP response to the deviant stimulus. In this difference wave, the MMN is observed as a negative peaking at 150-250 ms after the onset of the deviant stimulus. This component of negative polarity is typically detected over the frontocentral scalp sites, whereas a positive component is observed over the opposite side of the Sylvian fissure when the signal is referenced to a mastoid or a nose electrode (Näätänen, 2012).

The main MMN generator is suspected to be located bilaterally within the supratemporal plane creating the ‘supratemporal MMN’ (Baldeweg, Williams, & Gruzelier, 1999). This subcomponent of the MMN is presumed to relate to the process of change detection in perception (Näätänen, 2012). Furthermore, other sources, which contribute to the generation of an MMN, are suspected in the vicinity of the frontal cortex (Opitz, Rinne, Mecklinger, von Cramon, & Schroger, 2002). Presumably, the frontal subcomponent reflects the beginning of a spontaneous attention deflection provoked by the disruption of acoustic regularity (Escera, Alho, Winkler, & Näätänen, 1998).

The elicitation of an MMN occurs independently of the listener’s state of attention and thus is assumed to reflect preattentive processes (Alho et al., 1998; Näätänen, 2000; Näätänen & Michie, 1979; Näätänen, Teder, Alho, & Lavikainen, 1992), exemplified by the observation that an MMN is elicited in comatose adults (Luaute et al., 2005). Adult participants are often instructed to perform a distraction task in experimental situations (e.g., watching a film) to avoid overlapping of attention- or task-related ERP components with the MMN (Näätänen, 2012).

The rage of sound feature deviations that evoke an MMN is broad, encompassing intensity, duration, timber and even complex ones such as phonetic changes in speech sounds (Dehaene-Lambertz, 1997). The elicitation of the MMN is not only confined to auditory stimuli, but is also evoked by deviations in other modalities, such as the visual (Tales, Newton, Troscianko, & Butler, 1999) and somatosensory system (Kekoni et al., 1997).

The MMN was found to serve as an objective tool for capturing auditory discrimination processes. It was shown that the amplitude of the MMN increases with a higher grade of acoustic difference between the frequently presented and the deviant stimulus (Tiitinen, May, Reinikainen, & Näätänen, 1994). The discrimination threshold of acoustic deviance that elicits an MMN response corresponds approximately with the discrimination threshold observed in behavioral studies (Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004).

The neural mechanisms beyond the MMN response are currently considered to be a ‘comparison process’ between the neural trace formed by the infrequent (deviant) stimulus and a memory trace that has been formed by the preceding frequent stimuli (Cheour et al., 2000; Näätänen, 2012). These memory traces are suggested to represent sensory inferences computed from representations of acoustic regularity that are formed in response to the frequent stimulus. A stimulus that violates the preceding acoustic regularity consequently elicits an MMN response (Winkler et al., 1996). Deducing from the duration of the inter-stimulus-interval (ISI) that still leads to the generation of an MMN response, these traces are maintained for five to ten seconds in the auditory sensory memory (Näätänen & Picton, 1987; Sams, Hari, Rif, & Knuutila, 1993).

Recent studies suggest that long-term memory exerts an influence on the preattentive processes of sensory auditory memory. Changes in the MMN amplitude, latency or duration elicited by a deviant stimulus that was familiar to the participants, provide evidence of a connection between the memory systems. Therefore, MMN represents an objective measuring method not only to asses sensory memory processing, but also to indirectly detect the influence of long-term memory on these traces formed by the sensory memory (Atienza & Cantero, 2001; Beauchemin et al., 2006; Huotilainen, Kujala, & Alku, 2001; Näätänen et al., 1997).

Taking the observation made on the adult MMN into consideration, its infant counterpart, the MMR, might also be a promising objective tool to measure the formation of long-term memory in developing populations. The MMR might complement the behavioral methods that have been used to assess long-term memory formation in infancy in the past.

1.5.2 Mismatch response (MMR) in infancy

The deviance detecting response in infants differs in many aspects from the traditional MMN observed in adults. Some authors have titled the infant precursor as mismatch response (MMR) (Brannon, Libertus, Meck, & Woldorff, 2008; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002). The term MMR is also used to describe the infant auditory deviance detection response targeted in our study. This less defining name was picked since the reported polarity and subcomponents of the auditory deviance detection response in infants vary immensely between studies on the MMR.

The MMR is elicited in infants by a deviant stimulus differing in pitch from the standard stimulus (Alho, Sainio, Sajaniemi, Reinikainen, & Näätänen, 1990; Háden, Németh, Török, & Winkler, 2016; Leppanen et al., 2004; Martynova, Kirjavainen, & Cheour, 2003), by deviance in stimulus duration (Friederici, Friedrich, & Weber, 2002; Trainor et al., 2003) and by changes in speech sounds (Cheour-Luhtanen et al., 1996; Dehaene-Lambertz & Pena, 2001). The dynamics of the morphology of MMR components and its neural correlates throughout development are a controversial issue. Due to diverse study results, different theories on the development of acoustic deviance detection and its neural correlates exist (Jing & Benasich, 2006; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002).

The auditory MMR components are often compared to adult components. Transferring neural correlates of adult ERP components on infant ones should be done with caution since there are crucial differences in processing auditory deviance between the different age groups (Cheour, 2007; Cheour, Ceponiene, et al., 2002; Kushnerenko, Van den Bergh, & Winkler, 2013). However, infant MMR components have often been considered as counterparts of the adult ones and thus might provide hints of the underlying development in auditory deviance detection (Kushnerenko et al., 2013).

Though there is little known about the scalp distribution in infancy, differences between the adult and the infant auditory deviance detection regarding topography exist. One of them is the distinct scalp distribution of the infant MMR response. While the adult MMN response dominates in the frontocentral scalp sites, the infant MMR is mostly observed with a broader scalp distribution (Cheour, 2007). To gain a better understanding of the scalp distribution in different age groups, it should be considered when analyzing infant MMRs. This step is performed by grouping the responses across the targeted regions of interest (ROIs) (Hoehl & Wahl, 2012).

In a longitudinal study on the ontogeny of auditory ERPs, Jing and colleagues captured dynamics in the topography of the MMR. They observed that by the age of six months, an MMN-like negative response prevails parietally and subsequently moves steadily to the frontal areas at the age of two years (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002). Other studies showed that an infant MMN-like response is elicited with a slightly longer latency than the adult MMN (Cheour, Alho, et al., 1998; Cheour-Luhtanen et al., 1996).

Furthermore, the state of alertness influences MMN in adults differently than the infant MMR. Stable MMN-like negative responses were obtained in neonates in all stages of sleep (Cheour, Martynova, et al., 2002; Hirasawa, Kurihara, & Konishi, 2002; Martynova et al., 2003). There exist very few studies on the effects of sleep on the MMR in the first month of life. However, there are other evidences that sleep affects MMR components in two-month-old infants (Friederici et al., 2002; Otte et al., 2013). Altogether, the state of alertness on the auditory deviance detection response seems to play a bigger role in adults and might gain more influence on the MMR in the course of maturation of the central nervous system (Cheour, 2007).

To gain a better understanding of the auditory deviance detecting responses between the targeted age of two and four months, it is crucial to understand its development from birth.

1.5.2.1 MMR in the neonatal period

Regarding amplitude and polarity, auditory deviance detection responses go through changes during development. Some authors reported an early negative component at the beginning of life appearing similar to the classical adult MMN response (Alho et al., 1990; Cheour-Luhtanen et al., 1995; Hirasawa et al., 2002; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002; Leppanen et al., 2004).

This early negative MMN-like component is considered by some as a reflection of the most ancient discrimination process of the maturing brain (Cheour-Luhtanen et al., 1996; Näätänen, 2000). Corroborating the theory that the MMN emerges early in ontogeny, a similar negative response to a deviant acoustic stimulus peaking at 100-250 ms was obtained in preterm infants between thirty and 35 weeks of conceptional age (Cheour-Luhtanen et al., 1996) and its magnetic analog (called MMNm) was found in fetuses at the age between 35 and forty weeks of gestational age (Huotilainen et al., 2005). The presence of a response to acoustic deviance during prenatal life proposes that mechanisms for auditory discrimination are present from fetal life on and thus are assumed to play a fundamental role in ontogeny (Cheour-Luhtanen et al., 1996).

Kushnerenko et al. labeled the mentioned MMN-like early negative component as “early negativity [EN]” (2013, p. 3) and suggested that it might only be observed under specific test condition, since the EN seems to be elicited in neonates when the auditory deviance involves a large change in spectral width (Kushnerenko et al., 2013).

Besides the adult-like EN with its peak at 100-300 ms latency, another positive component peaking frontocentrally at about 300 ms post-stimulus is assumed to be part of the auditory deviance detection response in the neonatal period. This positive component (PC) has been observed in many studies in newborns (Dehaene-Lambertz & Pena, 2001; Kushnerenko et al., 2007; Leppanen et al., 2004). A third component is discussed to play a role in the neonatal auditory deviance detection process, a long-lasting late negativity (LN), peaking at 400ms after stimulus onset (Kushnerenko et al., 2013; Martynova et al., 2003).

The components mentioned above have been described as part of the deviance detection response and occur independently or together. Stable representations of all three components were elicited in neonates when the deviant stimuli were of large spectral width (such as environmental noises, e.g., bird chirps), thus suggesting that neonates require a wide spectral deviation for auditory discrimination (Kushnerenko et al., 2007).

1.5.2.2 Between two and six month

He and colleagues compared components of the auditory deviance detection response between different age groups. They found an EN response to pitch changes of piano tones in four-month-old, but not in two-month-old infants. However, this negative response in four-month-old infants was longer in latency but showed shorter latencies and increased amplitudes when widening the pitch difference between the deviant and standard stimulus. In contrast, two-month-old infants showed a broad PC to pitch deviance, but no stable EN. These results suggest that the resolution of pitch discrimination refines between the age of two and four months (He, Hotson, & Trainor, 2009).

Differing results concerning the time of appearance of the EN were provided by Jing and Benasich in a longitudinal study. They monthly obtained ERP data from five subjects between the age of three months and two years encompassing MMRs to pitch changes in an oddball sequences. According to their observations, a robust EN was observed between six and seven months with an increasing latency mainly in the first year of life (Jing & Benasich, 2006).

An explanatory approach to the absence of the EN in the first month of life is given by Kushnerenko and colleagues who suggested that other auditory discrimination ERP components might mask the classical EN at certain stages of development. In a longitudinal study, they investigated the development of the MMR to pitch changes in the first year of life. Particularity between the age of three and six months, the EN showed a tendency of being overlapped by the subsequent positive component. This PC was observed at birth and increased three-fold in the first three months of life. They found a stable differentiation in an EN-PC-LN complex (with peaks at 200, 300, and 450 ms) of the infant MMR at the age of six months (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002).

Looking at the diversity of components that have been reported and defined as MMR from birth to infancy, it raises more question of their origin. One theory postulates that the variance in results regarding the MMR polarity is caused by methodological differences between the studies, such as the type of oddball paradigm, sleep versus wakefulness, stimuli and filtering (Cheour, 2007; Hirasawa et al., 2002; Martynova et al., 2003; Trainor et al., 2003). Others suggest, that the EN and PC are separate components originating from distinct neuronal processes (He et al., 2009; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002) or that their appearance is associated with the age of the subjects (Kushnerenko et al., 2013). Furthermore, high interindividual differences in the ERP responses, as often observed in infants, might contribute to the diverging results (Trainor, 2007).

1.6 MMN as a tool for detecting voice familiarity

Only a few studies have investigated the influence of voice familiarity on the MMN in adulthood. A study conducted by Titova and Nätäänen provided evidence that the MMN can be used to detect neural processes that are involved in voice discrimination. Unfamiliar voice stimuli were presented to the participants as deviant and standard stimuli in an oddball paradigm. The deviant stimuli evoked a frontocentrally peaking MMN response whose mean amplitudes corresponded highly with behavioral dissimilarity ratings between the voices used as standard and the ones used as deviant stimuli. The study proposed further that processes involved in voice discrimination partly occur at a preattentive level (Titova & Naatanen, 2001).

Exploring electrophysiological processes of voice familiarity, the research team of Beauchemin found that the MMN and the P3a component are suitable tools for measuring the neural processing of voice recognition in adults. In their study, the presented oddball paradigm consisted of the vowel ‘a’ pronounced by different voices. A familiar deviant stimulus was provided by a friend of the participant and an unfamiliar deviant stimulus by an unfamiliar speaker. For analysis, the MMNs and the P3a components to both deviant stimuli were compared and their amplitudes showed significant differences between the responses. Through their observations Beauchemin and colleagues postulated that they had found an electrophysiological correlate for voice-specific brain areas that are tuned to familiar voices. From the observed influence of voice familiarity on preattentive ERP components, such as on the MMN, Beauchemin deduces that there is a “direct connection from long-term memory to the feature-analysis system of the short-term memory used to detect the acoustic features of stimuli” (2006, p. 3085).

In line with this theory, it was shown that familiarity with the deviant stimulus has an impact on the required number of standard stimuli to elicit an MMN. The familiarity with a deviant but native vowel decreases the number of required standard stimuli to evoke an MMN. These findings support the assumption that long-term memory representations influence on the formation of short-term memory traces (Huotilainen et al., 2001). Summing up, it has been suggested that voice familiarity affects the preattentive processes that lie behind the MMN. The observations made by Huotilainen and Beauchemin (2006; 2001) lead to the conclusion that preexisting long-term memory representations for familiar voices modulate the formation of short-term memory traces.

1.7 The influence of voice familiarity on ERP components in early childhood

As described before, voice familiarity in newborns and infants has been measured by behavioral methods (e.g., movements or sucking behavior) for decades (Moon, Zernzach, & Kuhl, 2015; Querleu et al., 1984). Behavioral studies provided cues for the existence of long-term memory representations for the mother's voice in early life. However, only a few studies explore electrophysiology behind these networks. While there is one adult study (Beauchemin et al., 2006) and just one published study in neonates (Beauchemin et al., 2011) that uses the auditory deviance detection response as a tool for investigating voice familiarity, no studies have been conducted to explore the long life span between birth and adulthood.

Cheour and colleagues showed in a longitudinal study that speech memory traces modify the infant EN amplitude (they used the term MMN). Native vowels elicited EN responses higher in amplitude than foreign vowels by the age of one year. These results might provide electrophysiological evidence that a connection between the long-term memory templates for speech sounds and the formation of short-term memory traces, might be present to some extent in infancy (Cheour, Ceponiene, et al., 1998).

For a better understanding of how full-term neonates (eight to 27 hours of age) process their mother's voice on a preattentive level, Beauchemin et al. compared the MMR to the mother's voice to the MMR elicited by a stranger's, both presented in an oddball paradigm consisting of vowels. Their results provided evidence that the two different voice stimuli induce different neural responses. A Source Analysis revealed that the mother's voice was preferentially processed in the left temporal lobe at early latency and subsequently activated right central regions, while the strangers' voices elicited mostly regions of the right temporal lobe. Taking these different activation patterns into consideration, Beauchemin and colleagues concluded that the mother's voice induces predominantly language-relevant processing whereas the stranger's voice evokes a voice-specific response. Furthermore, Beauchemin and colleagues deduced form their observations that tuning of certain brain areas to voice familiarities is apparently present from birth (2011).

Other studies investigated auditory discrimination and cognitive processes behind endogenous and exogenous auditory ERP components elicited by the mother's and a stranger's voice, assorted in voice paradigms. However, the assortment of these paradigms differed in the probability of stimulus occurrence from the oddball paradigm that was used in our study. These studies are outlined here, but their results should be differentiated from MMR studies since another type of paradigm was used, no difference waves were computed and other components were analyzed.

DeRegnier and colleagues investigated auditory recognition memory in newborns through voice familiarity. They used a voice sequence that contained the word ‘baby’ alternatingly pronounced by the mother and female stranger. The ERP component analyses showed differences in a positive peaking component (P2) and in a negative slow-wave component (NSW). The P2 component, peaking at 150-400 ms after stimulus onset, was found to be greater in amplitude and longer in latency to the maternal voice than to the unfamiliar voice. DeRegnier and colleagues interpreted that the differences in P2 amplitude reflect discriminative processes. The late NSW component was observed in response to the unfamiliar voice and was interpreted to reflect novelty detection (deRegnier, Nelson, Thomas, Wewerka, & Georgieff, 2000). The same study group showed that two weeks of postnatal experience changed the previously described P2 wave in response to the same paradigm; the early positive component was longer in latency and demonstrated more complex wave formation when elicited by the mother’s voice (deRegnier et al., 2002).

Another theory on how voice familiarity develops in the first month was provided by Mai and colleagues. They assumed that neonates have mechanisms for differentiating between “‘mother’s voice’ or ‘non-mother’s voice’” (2012, p. 3) that involve discriminating between voices irrespectively of voice familiarity. At an older age and after the exposure to a wide range of other voice stimuli, this mechanism might turn into a process similar to voice recognition; voices are categorized as the “‘mother’s voice (familiar) or ‘the other voice (novel)’” (Mai et al., 2012, p. 3). This approach implies the assumption that encoding of novel stimuli occurs at a later stage. Mai and colleagues conducted an ERP study in awake two-month-old infants based on a voice paradigm similar to the one used by deRegnier (2002). They made observations in respect of the P2 wave amplitude and the NSW amplitude consisted with the result found by DeRegnier in neonates (Mai et al., 2012).

Purhonen and colleagues demonstrated that the mother’s voice presented as a deviant stimulus in an oddball paradigm evoked early obligatory auditory ERP components shorter in latency than those elicited by an unfamiliar deviant voice in 4-month-old infants. On the contrary, later components that were suggested to indicate cognitive processing showed longer latencies for the mother’s voice than for the stranger’s. Purhonen and colleagues interpreted these findings also as a signal of pre-existing memory representations for the mother’s voice (Purhonen et al., 2005).

The formation of networks for voice recognition and stable long-term memory representations for the mother’s voice presumably develop prenatally and mature in the first months of life. However, very little is known about the maturing processes that lie behind the formation of these representations and their neuroanatomical correlates.

2 OBJECTIVES AND HYPOTHESES

The purpose of this study was to obtain and compare MMRs from healthy infants of the age between two and four months. The MMRs were elicited by two differing voice conditions: a familiar voice, represented by the mother's voice, versus an unfamiliar voice. The two voice stimuli were arranged as deviant stimuli besides a standard voice stimulus in an oddball paradigm and thus were presented in a constellation that targeted to elicit an auditory deviance detection response.

The MMR is described as the infant auditory deviance detection response and has successfully been used as a method for investigating preattentive processes in developing populations. Due to the non-invasiveness of EEG measures and the fact that attention is not needed to elicit an MMR (Cheour, 2007), it is a suitable ERP method for infants (deRegnier, 2007).

Behavioral studies showed that familiarity with the mother's voice is detected back to early stages of ontogeny. As described above, few electrophysiological studies in newborn and infants have been conducted investigating ERP responses to voice familiarity, only one used the MMR as targeted component (Beauchemin et al., 2011). Furthermore, we planned a pre-analysis of the original ERP responses elicited in response to the voice stimuli. In accordance with the study of Beauchemin and colleagues, a comparison of the original ERP responses to the standard stimuli with the responses to both deviant (familiar and unfamiliar) stimuli might clarify differences in the resulting MMRs (2011).

Neural processes that underlie mechanisms of voice recognition in infancy and their dynamics during ontogeny is not well understood. Adding one missing piece to a puzzle, we aimed to investigate the differences in the MMRs to a familiar as well as to an unfamiliar voice. We assumed that familiarity with the mother's voice, as a part of auditory recognition memory, might modulate the auditory deviance detection response on a preattentive level and thus could lead to differences in MMR characteristics. We searched to gain a better understanding of the processes behind voice familiarity, recognition and the formation of long-term voice memory representation in ontogeny.

Furthermore, we intended to find reliable electrophysiological correlates for voice familiarity in infancy and range our results into the preexisting observations by other researchers. To enable a comparison, we adapted the oddball paradigm to paradigms used in ERP studies on voice familiarity in newborns (Beauchemin et al., 2011; deRegnier et al., 2000; deRegnier et al., 2002).

For longitudinal notion of how characteristics of the MMR to voice familiarity behave throughout different age stages, it is crucial that MMR studies are conducted in different age groups. We chose the targeted age group to provide information on a stage of development at which the infant already had been exposed to a high amount of varying voices and might have formed a stable long-term memory representation of the mother’s voice. Furthermore, the targeted age might coincide with a milestone in the ontogeny of auditory processing; enhanced synaptogenesis within the auditory cortex is observed at the age of three months (Huttenlocher & Dabholkar, 1997), reflecting a neuroanatomical substrate that is involved in the formation of auditory recognition memory and voice recognition in adults.

3 METHODS

The present study was embedded into a main project that investigated sleep-dependent memory consolidation of voice recognition in early infancy. The data we captured for the main project were partly used in our study. In the study run for the main project, the initial presentation of the oddball sequence was followed by a phase of familiarization (duration: 10 min) that encompassed the presentation of the unfamiliar voice previously used as an unfamiliar deviant stimulus in the voice paradigm. Subsequently, a retention period succeeded consisting of wakefulness or/and sleep phases. Polysomnography was recorded during that period to assess sleep stages and sleep timing (duration: 90-150 min). In a following recall session, a similar version of the initially used voice paradigm was re-presented and ERPs were recorded. The paradigm used in the recall session differed only in the voice that provided the standard stimulus from the initially presented paradigm.

The procedures and data regarding the sleep-related condition would go beyond the scope of the following descriptions and will not be specified. The present study solely focused on the first part of the described study run, the ERPs elicited by the first presentation of an oddball voice paradigm.

3.1 Recruitment and participants

17 female and 18 male healthy, full-term infants between the age of ten and 16 weeks (*M*: 98.5 days, *SD*: 17.4 days, range: 53-103 days) accompanied by their mothers participated in the study. The study obtained approval of the ethics committee of the University Hospital Tübingen and its Faculty of Medicine (number of vote: 512/2013BO1).

Before participating and after receiving information on the procedure, the mothers gave written informed consent for their children. An expense allowance was provided for the participating families encompassing 15 Euro per hour. The recruitment was achieved via the university's email distribution list, former studies and advertisements (e.g., flyers) placed at public places of the university and childcare facilities. In preceding telephone conversations and email contact, the participants' parents were informed about the study procedure following a standardized telephone protocol. Eligibility was surveyed during the telephone call and again on the day of the testing using a standardized questionnaire. Information of one participant was not inquired since no ERPs were recorded.

Prerequisites for a participation were singleton full-term birth (*M*: 39.8 weeks, *SD*: 0.95 weeks), while the birth modus was not taken into consideration. Participants with congenital malformations or low birth weight (< 2500 g) were not included. Further exclusion criteria were severe complications during pregnancy or delivery of both, the mother and the child (e.g., not dietetically adjustable gestational diabetes, cardiopulmonary resuscitation), known neurological illnesses or abnormal results in either the newborn auditory screening or follow-up screening tests.

Complications during pregnancy were assessed. Four of the mothers had had a diagnosed gestational diabetes that had been adjusted dietetically, two reported an iron deficiency during pregnancy, one had had a thyroid deficiency that had been diagnosed during pregnancy. Another mother referred to gestational hypertension that had been treated with methyldopa. For assessing complications during postnatal adaption, the participants' Apgar score was inquired (Apgar, 1953). All participants showed an Apgar score greater than or equal to 6 after one minute (*M:* 8.65, *SD*: 0.92), a score greater than or equal to 8 after five minutes (*M*: 9.33 *SD*: 0.65) and a score greater than or equal to 9 after ten minutes (*M*: 9.8 *SD*: 0.45).

Furthermore, body dimensions at the time of birth were collected via the pregnancy record booklet (*Mutterpass*) and percentiles were determined: head circumference (*M*: 35.4 cm, *SD*: 1.33 cm; *M*: 57. percentile, *SD*: 25.3), birth weight (*M*: 3520 g, *SD*: 457.82 g; *M*: 48. percentile, *SD*: 27.31), birth length (*M:* 51.6 cm, *SD*: 2.2 cm; *M*: 59. percentile, *SD*: 27.31) (Fenton & Kim, 2013). Two infants were born as small for gestational age (<10. percentile), two were born as large for gestational age (> 95. percentile) and 31 were born appropriate for gestational age. All interviewed mothers stated to breastfeed their child (n=34), some of them partially added formula to the diet (*n*=6).

Information on the mothers was collected encompassing age (*M*: 30.7 yr, *SD*: 5.38 yr), the number of births (*M*: 1.7, *SD*: 0.95) and consumption of tobacco (*n*=1). Some mothers stated occasional moderate alcohol consumption (*n*=8) during lactation. Six mothers received hormonal treatment with levothyroxine due to thyroid hormone deficiency. Others took vitamin or micronutrient preparations (e.g., folic acid, vitamin B12, iodine, zinc) (*n*=6).

The mothers were further asked about their highest level of school education: the majority indicated general qualification for university entrance (*Abitur*) (*n*=24). The languages spoken at home were inquired. Children from families that spoke other languages than (*n*=1) or beside (*n*=7) German at home were included in the study.

Furthermore, sleep habits including the sleeping time the day of the testing, longest sleep interval, quantity, preferred place (bed, cradle, arm etc.) and sleep position were collected via an anamnesis questionnaire. For the main project, 15 mothers filled out the Brief Infant Sleep Questionnaire (BISQ), that assesses more detailed information on sleeping behavior in early infancy (Sadeh, 2004). For estimating the participant's state of alertness on the day of the testing, the mother was asked about her child's sleeping amount (*M*: 13.5 h, *SD*: 2.02 h) and variations from daily routine in the last 24 hours. Two children had received vaccination two days before the testing.

While scheduling the testing time, the baby's sleeping behavior was taken carefully into consideration. We targeted that the presentation of the paradigm was performed in a period of alertness. One test session was not completed because the infant did not tolerate the EEG-cap and cried. Further EEG-data of two infants were excluded from the initial sample for further analyses because the number of trials did not reach inclusion criteria or due to bad signal in EEG-recording.

3.2 Procedure and setting

The study took place in the premises of the children's sleeping laboratory in the Institute of Medical Psychology and Behavioral Neurobiology in Tübingen that provides a study suitable setting for infants (equipped with a changing table, cuddly toys etc.). As our project was integrated into another study, one session encompassed between three and five hours, varying in time due to the infant's sleeping and feeding behavior.

The mothers played a crucial part in the study setting; they were present all along the study run to support the investigator by ensuring the child's individual needs (e.g., to hold, cuddle and breastfeed the baby). Moreover, the mothers had an active part in the study since their voices were recorded at the beginning for generating a stimulus for the subsequent voice paradigm. A study run was conducted following a standardized study protocol. Before presenting the oddball voice paradigms to the infants, the mothers were asked to estimate if their child needed breastfeeding or diapering beforehand. That way the infants' needs were satisfied before we started the study run.

3.3 Voice oddball paradigm

An oddball voice paradigm was presented to the participants that targeted to elicit ERPs as brain responses to the different stimuli. The paradigm's assortment was adopted from previous studies using a similar voice paradigm for ERP elicitation during the first year of life (Beauchemin et al., 2011; deRegnier et al., 2000; Mai et al., 2012). It consisted of three different female voice stimuli providing three different conditions: (1) a frequently presented 'standard stimulus' provided by an unfamiliar voice (85% of all stimuli), (2) the mother's voice as an infrequently presented 'familiar deviant stimulus' (7.5%) and (3) another infrequently presented deviant stimulus vocalized by an unfamiliar voice (7.5 %), here referred to as 'unfamiliar deviant stimulus'. The stimuli were arranged in a repetitive sequence of the word 'baby' (deRegnier et al., 2000).

One trial consisted of one voice stimulus (400 ms) and the interstimulus interval (ISI), the time between two stimuli, lasted 600 ms. Consequently, a stimulus was presented every 1000 ms. A complete run through the paradigm took ten minutes which leads to 600 as the total number of trials per study run. The sequence was composed of a fixed proportion of the three different stimuli following the scheme of the oddball paradigm used in the study about voice processing in newborn infants by Beauchemin and colleagues (Beauchemin et al., 2011).

The familiar deviant stimulus, was provided by the mother's voice and was recorded in advance on the day of the testing (see Voice recordings). The stimuli were arranged in a pseudorandomized order with the proviso that every infrequent stimulus was followed by at least three standard stimuli. This approach was chosen to increase the chance that a neural trace for the standard stimulus was formed (Beauchemin et al., 2011). For the presentation of the oddball paradigm the software Presentation® (Neurobehavioral Systems, Berkeley, CA, United States) was used. The unfamiliar stimuli were chosen of a pool of unfamiliar voices encompassing two different female voices for the unfamiliar deviant stimuli and four different voices for the standard stimulus respectively. The allocation of the unfamiliar voices to the stimuli varied between the study runs and accorded to a balanced scheme.

During EEG recording, the paradigm was constantly presented binaurally through loudspeakers with a sound pressure peaking at 75 dB to avoid an ERP elicitation caused by deviance in intensity between the stimuli. The volume peak was measured with a volume level-measuring device in advance. The two loudspeakers were positioned on both sides of a changing table at the babies' auricular height. For the voice presentation, the babies were placed on it, in a way that the distance from each ear to each loudspeaker was about 45 cm (see Figure 3.1).

The voice paradigm was presented while the children were awake. 15 infants had slept and most of them had been fed (*n*=28) right before the voice presentation. To assess the children's state of alertness during presentation 23 of the mothers were asked previously to estimate their child's sleepiness by using a scale from 1 ('wide awake') to 10 ('fast asleep'). All infants were estimated smaller than or equal to 6 (*M:* 2.86, *SD*: *1.6*).

All along the voice presentation, the mothers were standing in front of the baby changing unit in order to distract their child visually by playing with hand puppets or grimacing (Brannon et al., 2008). This approach was chosen to provide proximity of the mother and to keep the infant attentive (Hoehl & Wahl, 2012). In advance, the mothers received instructions to make no sounds or talk during the voice presentation to minimize acoustic stimulation other than the paradigm. The interaction between mothers and children as well as infants' effusive movements (e.g., kicking and sucking on a pacifier) were documented during the presentation of the paradigm. The experimenter was located in an adjoining room entering once in a while to make observations or, if necessary, assist the mothers in distracting their children. Moreover, the experimenter had access to the computer that generated the paradigm and could pause the paradigm in case the child got agitated or started to cry *(n=2)*.

Figure 3.1: Participant in the study setting.

An Infant with an EEG cap placed on the baby changing table in the study setting.

3.4 Voice recordings

For the oddball paradigm voice recordings of both the mother's and the unfamiliar female voices were required. All voice recordings were recorded via a portable USB Condensor Microphone (Go Mic® by Samson Technologies) with a handmade pop filter. For recording and editing the voices, the program Audacity® 2.0.5 was utilized.

Four women between the age of twenty and thirty years provided their voices as unfamiliar stimuli for the voice paradigm. These stimuli were recorded in advance in the same room. The mothers' voices were recorded and edited on the day of the testing. Before the voice recordings were realized, the experimenter gave the mothers instructions on how to pronounce the word 'baby'. The women were guided to speak the word ‘baby’ with a German pronunciation and emphasis on the first syllable (phonetics: [ˈ*beːbi*]) for several times while keeping their voices as natural as possible and avoiding emotional connotation. The stimuli were not recorded as an infant-directed speech. Subsequently, a suitable stimulus was selected among the recorded sequence by taking pronunciation and recorded stimuli length (possibly close to the stimulus duration of 400 ms) into account.

All stimuli were edited by carrying out the same steps to create stimuli that were standardized in length and quality: the effects ‘Noise Removal’, ‘Amplify’, ‘Change Tempo’ in Audacity® were applied to the recordings.

3.5 EEG recordings

At the beginning, the subject’s head circumference was measured (*M*: 41 cm, *SD*: 1.32 cm) and an infant-suitable EEG recording cap (EasyCap®, by EASYCAP GmbH, Herrsching, Germany) in the appropriate size was chosen. The EEG-cap was either fixed with a chest strap or fastened with a strap underneath the child’s chin, while the mother distracted the child.

A passive electrode system arranged as a low-density montage was used to record the EEG-signal. For that, twelve Ag/AgCl sintered ring electrodes (EASYCAP GmbH) were fixed to the EEG-cap, filled with conductive gel and subsequently covered with medical adhesive tape. The gel was warmed up in warm water to make the application more comfortable for the child and ensure better contact of the electrodes to the scalp (Hoehl & Wahl, 2012).

For obtaining EEG recordings, the scalp electrodes F3, F4, Fz, FCz, C3, C4, Cz, Pz, M1 were referenced to the right mastoid electrode (M2) according to the international 10-20 system for scalp electrode position (Klem, Luders, Jasper, & Elger, 1999). A ground electrode was attached to the forehead at the FP2 position. The EEG signal was amplified by using a BrainAmps® DC amplifier (Brain Products GmbH Gilching, Germany) that was interfaced with a personal computer. The EEG signal was continuously recorded at an A/D conversion rate of 500 Hz while a band-pass filter of 0,1 - 250 Hz was applied.

Eye movements were recorded with Electrooculography (EOG) for subsequent electro-ocular artifact removal from the recorded EEG data. For that, one

EOG-electrode was integrated into the cap at the FP1 position; another was attached with a piece of medical adhesive tape and conductive paste below the left eye (EOG1) to trace vertical eye movements.

Electrode impedances were checked before every new ERP recording and attempted to be kept under 20 kΩ, adhering to the recommended range of the amplifier. The ground and reference electrodes' impedances were kept under 10 kΩ. The incoming signals were visualized and recorded by the program VisionRecorder® (Brain Products GmbH Gilching, Germany) on a PC that was connected to another PC generating the paradigm. When a stimulus was presented, a trigger signal was sent from the PC that generated the paradigm to the PC that recorded the EEG signals. That way, stimulus marker and the incoming signals were visualized simultaneously via VisionRecorder®.

3.6 ERP data editing

The EEG-data were pre-processed offline using the program BrainVision Analyzer 2 (Version 2.1.0 by Brain Products). The process of signal transformation was adapted to indications proposed in papers and literature about ERP techniques in infants (DeBoer, Scott, & Nelson, 2007; Hoehl & Wahl, 2012). First, the EEG signals were re-referenced to a channel that was formed of the average signals from the left and right mastoid electrodes (linked mastoids, M1, M2). When the M1 impedance was high or the signal contained many artifacts, this step was skipped for every condition and M2 was kept as the reference electrode (*n*=2).

3.6.1 Filtering and segmentation

A digital filter with a high-pass filter (half-amplitude cutoff: 1 Hz, roll-off: 12 dB/octave) and a low-pass filter (half-amplitude cutoff: 30 Hz, roll-off: 24 dB/octave), as well as a notch filter of 50 Hz were set. Segmentation of the continuous EEG signal was realized as followed: choosing a pre-stimulus interval of 200 ms and a post-stimulus interval of 1000 ms resulted in a total epoch length of 1200 ms.

One paradigm run produced in total 600 epochs consisting of 510 trials that contained the ERP responses to the standard stimulus as well as 45 epochs each in response to the familiar and unfamiliar deviant stimulus, respectively. Baseline correction was performed by applying the average voltage at the time of 200 ms before stimulus onset for each segment.

3.6.2 Artifact rejection

A new channel for vertical eye movements was calculated from the signals of the EOG1 and the FP1 channel by linear derivation. Blink artifacts were detected semi-automatically. Independent Component Analysis (Hoffmann & Falkenstein, 2008) was performed to detect artifacts followed by a visual inspection. Due to a great number of artifacts in the newly built EOG channel, Independent Component Analysis was left out for some subjects (*n*=3). Other artifacts were also detected semi-automatically.

Flatlines (indicating amplifier's saturation, blocking) of under 0.5 μV and gradients (voltage steps) exceeding 75 μV/ms, as well as the interval of 200 ms before and after the artifact were detected and classified as an artifact in the raw data of all channels. Furthermore, amplitude steps in the trials that crossed 100 μV within an interval of 300 ms, as well as absolute amplitudes of more than 100 μV and less than -100 μV, were detected as artifacts. Epochs containing one of the artifacts mentioned above were withdrawn from data analysis. Epochs with artifacts in the interval of -200-100 ms before stimulus onset, as well as trials with artifacts in the interval 800-1000 ms after stimulus ending were retained in the analyses. Independent Channel Rejection (ICR) was applied to proceed artifact rejection (Fujioka, Mourad, He, & Trainor, 2011): for each epoch that contained artifacts, all channels with detected artifacts were rejected, the others were kept for averaging.

For all subjects, the signal of the Pz electrode was excluded since it showed consequently a bad signal in most of the participants. We hypothesized that the signal of the Pz electrode was disturbed since the babies were lying in a supine position during the testing session. That way, the Pz electrode might have had contact with the surface of the baby-changing unit what might result in a disturbed signal.

3.6.3 Averaging and calculating difference waves

Electrode channels were excluded from further averaging when their total number of segments fell short of fifty for the standard condition and ten for the unfamiliar or the familiar condition, respectively. The criterion of the maximal number of epochs for the process of averaging was based on indications provided by publications on the ERP technique in infants (DeBoer et al., 2007). The mean number of included epochs per condition across all electrodes was 268.43 (52.2%) for the standard condition, 25.75 (57.23%) for the familiar condition and 25.59 (56.88%) for the unfamiliar condition (see Table 3.1 for the number of epochs per electrode position). The number of epochs that were included in the analysis of the familiar and the unfamiliar condition did not differ significantly from each other [$t(31)=-.191$, $p= .850$].

After excluding the data of one subject due to an insufficient number of trials and another due to bad signal over all electrodes, we obtained electrode-specific epochs from 32 subjects. Subsequently, signal averaging was performed resulting in an averaged wave for every condition per subject. For achieving the MMR, difference waves were calculated by subtracting the averaged standard wave from the averaged (familiar or unfamiliar, respectively) deviant wave.

Table 3.1: Mean number of epochs.

The mean number of epochs (*M)* per electrode site, its standard deviation (*SD*) and the number of participants (*n*) that contributed to the channel grand-average ERPs.

	Standard			Familiar deviant			**Unfamiliar deviant**		
	M	*SD*	*n*	*M*	*SD*	*n*	M	SD	n
F3	308	74,4	30	27,9	7,62	30	27,4	7,73	30
Fz	285	90,8	30	25,7	8,51	30	25,8	8,22	29
FCz	269	112,	30	25,5	8,72	27	26,1	8,08	26
F4	283	94,2	31	25,8	8,39	30	25,7	8,36	30
C3	267	95,6	31	25,5	8,29	29	25,3	8,35	29
Cz	254	105,7	29	24,6	9,18	26	25,0	8,46	26
C4	279	87,9	32	25,3	7,30	31	25,3	8,77	32
All	278	96,4	32	25,8	8,22	32	25,8	8,32	32

3.6.4 Quantifying ERP data

The step of quantifying ERP components implies defining quantitative criteria for their latencies and amplitudes. Subsequently, the resulting values are used to perform statistical tests. Quantifying is normally performed within a pre-defined measurement window (Luck, 2014b, Chap. 9). Adhering to specifications and results from preceding studies, we chose one measurement window of 100-300 ms and another in a time window of 300-450 ms after stimulus onset for the analysis of the original ERP responses as well as the MMRs.

For quantifying the amplitudes of the original ERP responses and the resulting MMRs, an area amplitude approach was chosen. The area amplitude (in ms·µV) is defined as the integral areas under the curve, signed with either positive values when located above the baseline and negative values when located under the baseline, were determined for the original and MMR waves in both measurement windows (Luck, 2014b, Chap. 9).

For analysis of the original ERP elicited by the three different conditions, signals were averaged in as frontal cluster (F3, Fz, F4) and a central cluster (C3, Cz, C4) for each time window. The inclusion of grouped scalp sites in the analysis of the original ERP responses was chosen following the study of Mai and colleagues (2012).

For analyzing the MMRs, a region of interest (ROI) was created by grouping the signal at the frontocentral electrodes (F3, Fz, FCz, F4, C3, Cz, C4) for each measurement window (Luck, 2014b, Chap. 10), here referred to as the frontocentral ROI.

3.7 Statistical analysis

For statistical analysis, the program IBM SPSS Statistics® was used. The assumption of normality for the area amplitudes in both measurement windows for the two different conditions, the age in days, the sleep in the last 24 h and the sleepiness score scaled by the mothers before the presentation of the voice paradigm (see section Voice Paradigm) were tested for normality with Kolmogorov-Smirnov tests.

Since our study was arranged as a within-subject design, repeated analysis of variance (ANOVA) was chosen as statistical model. To analyze differences between the area amplitudes of the original ERP responses to the standard condition and the area amplitude elicited by the two deviant conditions, 2x3 repeated ANOVAs were performed. The ANOVAs included the within-subject factors scalp site (‘frontal’ vs. ‘central’) and condition (‘standard’ vs. ‘familiar/ unfamiliar deviant’) for the two measurement windows separately. When the assumption of sphericity had been violated in the Mauchly’s test, a Greenhouse-Geisser correction was applied.

For testing the main experimental effect of voice familiarity on the area amplitude of the MMRs in the frontocentral ROI, repeated one-way ANOVAs were performed including the within-subject factor familiarity (‘unfamiliar deviant’ vs. ‘familiar deviant’) separately in the two targeted measurement windows. Since the within-subject factor solely contained two levels, no corrections for sphericity were needed. For an explorative analysis approach, separate *t*-tests on the area amplitudes of the MMRs at each electrode were conducted.

For visualizing the significant time windows on a finer grained level, two-tailed *t*-tests (Excel 2016 ®) at a sampling rate of two ms were performed on the signal of the grand-average MMRs. This procedure was carried out for each electrode separately as well as for the frontocentral ROI, resulting in 600 *t*-tests per epoch. The resulting *p*-values were plotted against latency as probabilities for accepting the null hypothesis for every time step.

Two-tailed Pearson’s correlation coefficients were calculated to explore the relationship between the possible confounders and the area amplitudes of the MMR to both deviant conditions. Correlations were tested between the age in days, the sleep in the last 24 hours as well as the sleepiness score inquired before presenting the paradigm and the area amplitudes of the MMRs to both deviant conditions assessed at the frontocentral ROI. When the assumption of normality for a variable was violated, a two-tailed Spearman’s correlation coefficient was computed.

Partial η^2 was reported to point out the effect sizes. The null hypothesis was rejected when $p< .05$. The results are reported as means and their standard error of mean (*M*±*SEM*).

4 RESULTS

4.1 Comparison of the original ERP responses

4.1.1 Visual comparison

Figure 4.1 shows the grand-averages of the original event-related potential (ERP) responses elicited by the standard, the familiar deviant (mother's voice) and the unfamiliar deviant (stranger's voice) conditions. Looking at the early measurement window, an early negative deflection, an N1-like peak (here labeled as N1), is discernible, followed by a positive deflection (here labeled as P2) at all fronto-central electrode sites for all conditions. With a closer look at the grand-average ERP responses in the early measurement window, a lower amplitude in response to the unfamiliar deviant condition (including N1 and P2 amplitudes) in comparison to the ERP responses to the two other conditions come to light.

Descriptively, the response to the standard stimulus differed most clearly from the response to the familiar deviant stimulus in the late measurement window, due to a second positive peak (here labeled as P3) that followed the P2. This additional positive deflection was distinguishable at all frontal and central electrodes, except at Cz. A similar positive deflection, however, lower in amplitude, seems to be elicited by the unfamiliar deviant condition in the late time window.

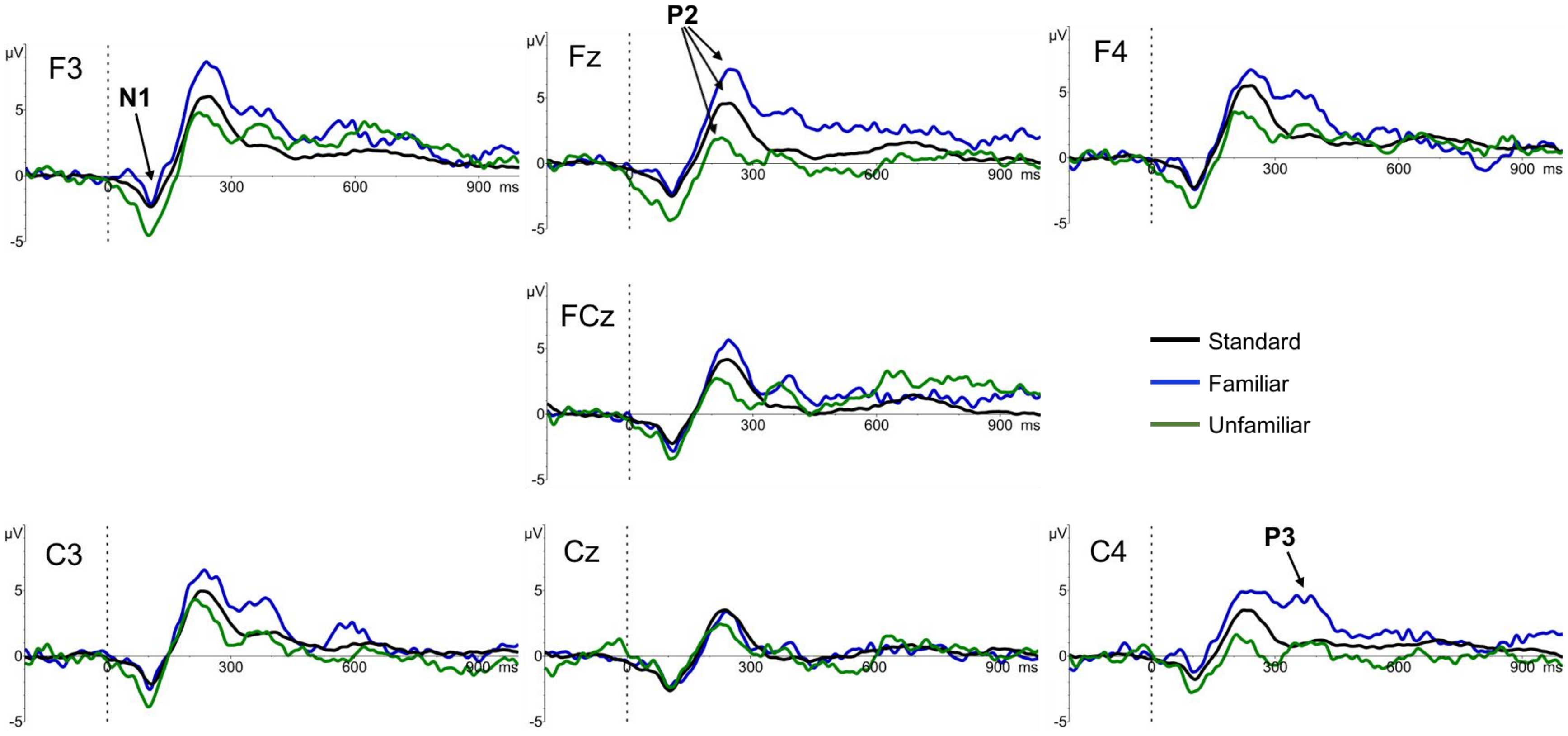

Figure 4.1: Original ERP responses.

The original ERP responses to the standard (black), the familiar (blue) and the unfamiliar (green) conditions shown at the frontocentral electrode sites. At all electrode sites, a negative (N1) and a positive (P2) component were distinguishable in the early time window (100-300 ms). Amplitudes of the P2 varied between the electrode sites. In the late time window (300-450 ms), a positive deflection (P3) was distinguishable in the ERP response to the familiar condition as well as to the unfamiliar condition at all electrode sites, except at Cz.

4.1.2 Statistical comparison

4.1.2.1 Early measurement window (100-300 ms)

For the early measurement window, the condition significantly affected the area amplitude [$F(1.44,44.7) = 4.76$, $p = .022$, $\eta_P^2 = .233$, $\varepsilon = 0.721$, Greenhouse-Geisser corrected]. Contrast revealed that the area amplitude in response to the unfamiliar deviant condition was significantly lower than the area amplitude in response to the standard condition [$F(31,1) = 5.7$, $p = .023$, $\eta_P^2 = .155$]. The mean area amplitude in response to the unfamiliar deviant stimulus was more negative at both the frontal (128 ± 142 ms·µV) and the central (36.6 ± 151 ms·µV) scalp sites than in response to the standard stimulus (frontal: 485 ± 91.3 ms·µV, central: 306 ± 86.1 ms·µV), as it is shown in Figure 4.2. The area amplitude in response to the standard condition did not differ significantly from the area amplitude in response to the familiar condition.

Furthermore, the scalp site reported to have a significant influence on the area amplitude [$F(31,1) = 11.78$, $p = .002$, $\eta_P^2 = .275$]. No interactions between the scalp site and the condition reported as significant, indicating that the effect of the condition on the area amplitude did not depend significantly on the scalp site at which it was measured.

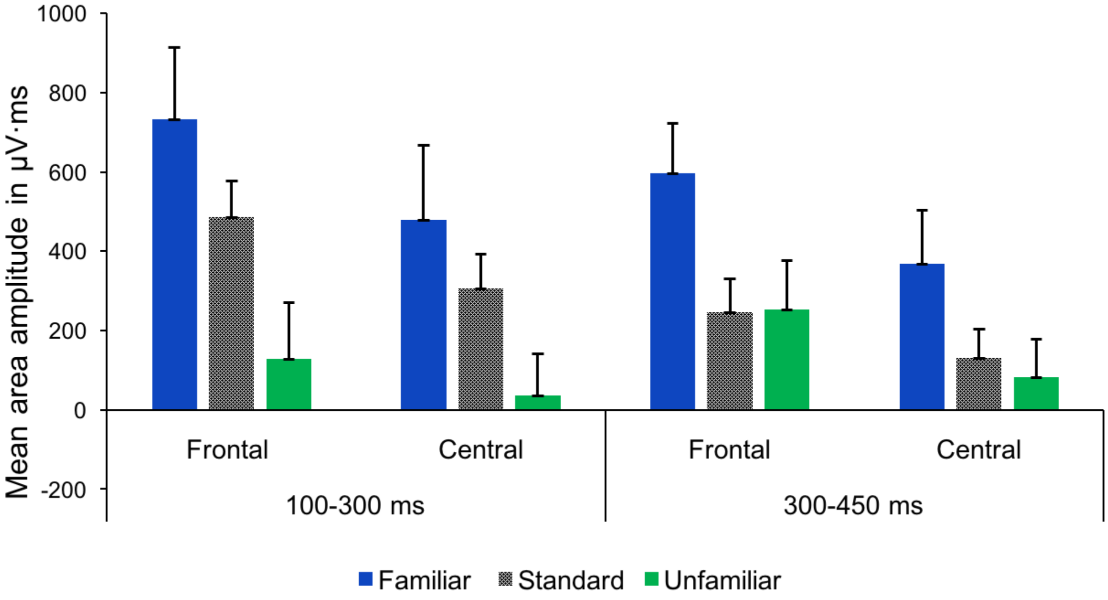

Figure 4.2: Mean area amplitudes of the original event-related potential (ERP) responses.

The mean area amplitude of the original ERP responses in the measurement windows of 100-300 ms and 300-450 ms after stimulus onset, shown for the frontal and the central scalp sites separately. The mean area amplitude of the original ERP responses to the familiar deviant (blue), standard (black) and the unfamiliar deviant (green) condition are contrasted. Error bars represent the standard error of the mean (SEM).

4.1.2.2 Late measurement window (300-450 ms)

For the late measurement window, the significant main effect of the condition on the area amplitude was revealed [$F(1.59, 49.4) = 4.84$, $p = .018$, $\eta_P^2 = .14$, $\varepsilon = 0.80$, Greenhouse-Geisser corrected]. The mean area amplitude in response to the familiar deviant condition at both frontal (596 ± 128 ms·µV) and central (387 ± 136 ms·µV) scalp sites showed to be more positive than in response to the standard condition (frontal: 247 ± 83.1 ms·µV, central: 131 ± 73.6 ms·µV), as it is illustrated in Figure 4.2. Contrasts revealed that the area amplitude was significantly higher in response to the familiar deviant than to the standard condition in the late time window [$F(1,31) = 9.99$, $p = .003$, $\eta_P^2 = .244$], as it is illustrated in Figure 4.2.

No significant difference was found between the area amplitudes in response to the standard and the area amplitude in response to the unfamiliar deviant condition in the late measurement window. Besides, the scalp site had a significant influence on the area amplitude [$F(1,31) = 19.6$, $p < .001$, $\eta_P^2 = .387$]. The area amplitude was significantly higher at the frontal scalp sites than at the central scalp sites, as it is shown in Figure 4.2. An interaction between the scalp site and the condition was not found to be significant.

Though significant differences in the area amplitudes were found between the central and frontal scalp sites for both measurement windows, a significant effect of the condition on the area amplitudes was observed at both scalp sites. Thus, we chose to analyze the mismatch responses (MMRs) at a frontocentral region of interest (ROI) that was calculated by averaging the signal of all frontocentral electrode sites.

4.2 Comparison of the mismatch responses

4.2.1 Multiple t-tests

Figure 4.3 illustrates the *p*-values of the multiple two-tailed *t*-tests and the corresponding MMR waves to the familiar and unfamiliar deviant condition at the single electrode sites. The multiple *t*-tests revealed significant differences in the amplitudes of the MMRs to the familiar deviant and unfamiliar deviant condition at F3 in a broad time window (232-296 ms post-stimulus) and several small time windows below the length of 20 ms. Significant differences in amplitude of the MMRs to the two deviant conditions were also found for the latency spans of 202-482 ms and 496-632 ms post-stimulus at Fz. As shown in Figure 4.3, the MMRs tended to be different in the time window between 150-330 ms post-stimulus at F4, but this difference did not reach significance. The amplitudes of the MMRs to the two deviant conditions in a time window of 334-420 ms post-stimulus significantly differed at C4. No appreciable statistically significant differences were found between the amplitudes of the MMRs to the familiar and unfamiliar condition at FCz, C3 and Cz.

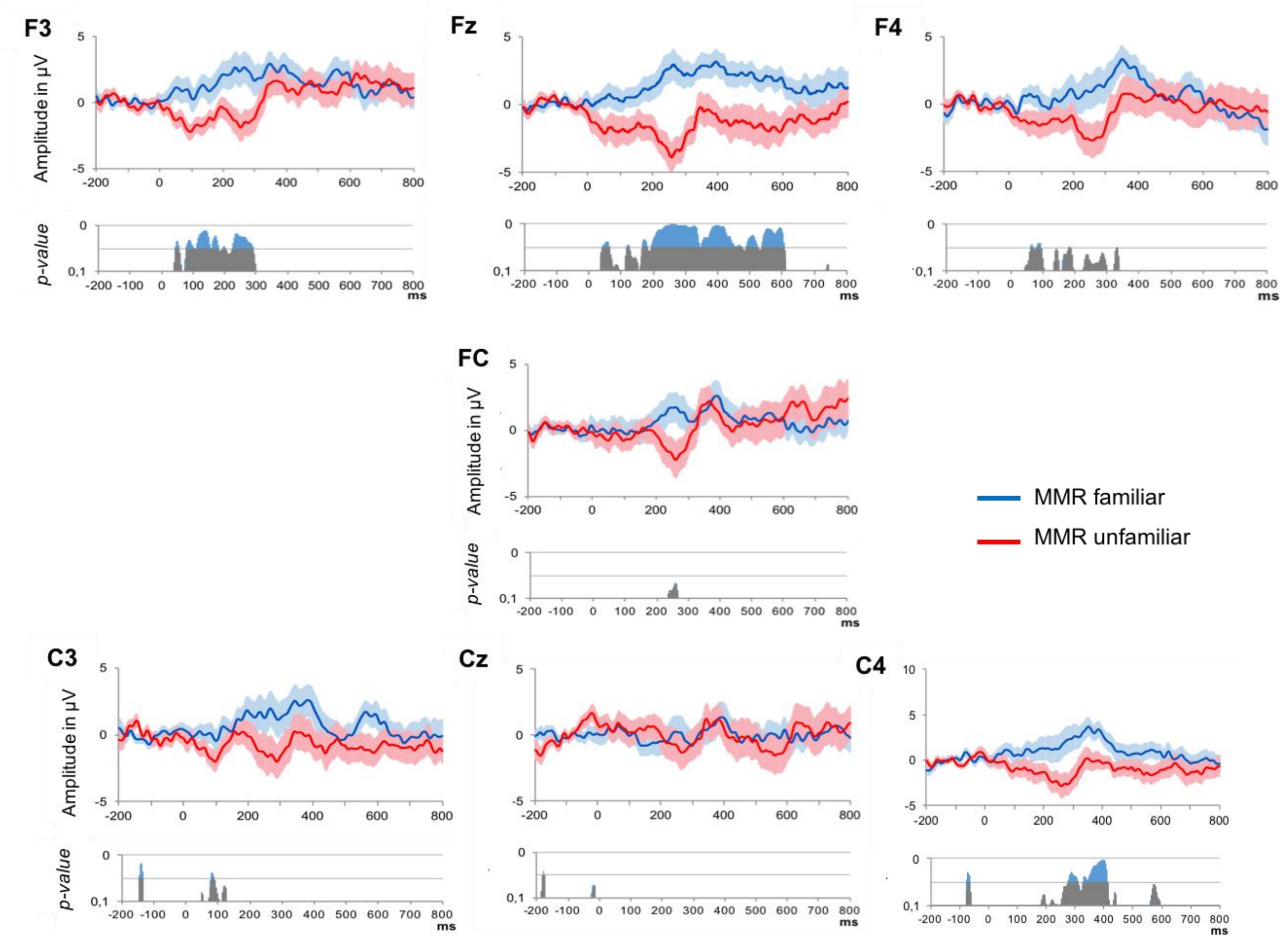

Figure 4.3: Mismatch response waves and multiple t-tests.

Mismatch response waves to the familiar (blue) and the unfamiliar (red) deviant stimulus at the frontocentral electrode sites, shown with their standard error of the mean (SEM) as broadened waveforms. Underneath the waves, the *p*-values of the multiple *t*-tests against time are depicted. *P*-values <.05 are highlighted in blue.

4.2.2 Visual comparison

The grand-average MMRs in the frontocentral ROI showed a more positive wave progression in response to the familiar deviant than for the unfamiliar deviant condition in both measurement windows, as it is illustrated in Figure 4.6.

The grand-average MMR to the unfamiliar deviant condition showed two negative deflections in the early measurement window (100-300ms) and a positive deflection in the late window (300-450 ms). On the contrary, a positive peak prevailed the grand-average MMR elicited by the mother's voice in the early measurement window, followed by another positive deflection in the late measurement window. The multiple t-tests revealed a significant time window approximately between 100 and 300 ms.

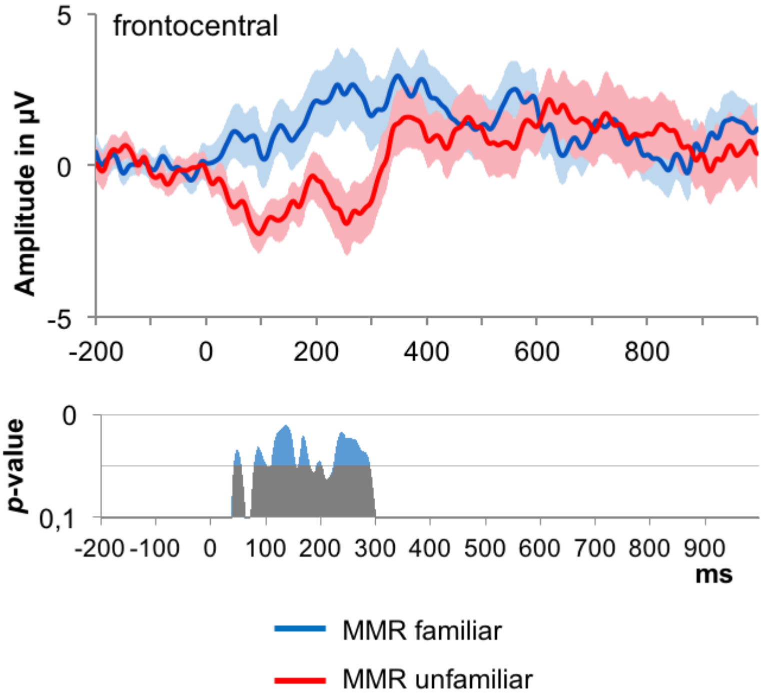

Figure 4.4: Mismatch response in the frontocentral region on interest

Mismatch response (MMR) waves to the familiar (blue) and the unfamiliar (red) deviant condition in the frontocentral ROI, depictured with the standard error of the mean (SEM) indicated as broadened waveforms. The MMR to the familiar deviant was more positive than the MMR to the unfamiliar deviant condition in the two targeted measurement windows (100-300 ms and 300- 450 ms). In the MMR to the unfamiliar deviant condition, two negative deflections are distinguishable in the early measurement window. Underneath the curves, *p*-values of the multiple *t*-tests against time are depicted. *P*-values <.05 are highlighted in blue.

However, visual inspection of the individual MMRs in the frontocentral ROI revealed a high interindividual variability in wave progression, reflected as differences in components, latency and polarity. Representative MMRs of two participants (1 and 2) that are comparable in the mean number of trials for the familiar deviant (M_1: 40, M_2: 41) and the unfamiliar deviant condition (M_1: 39, $M2$: 40) are illustrated in Figure 4.5. However, both examples show noise in the prestimulus baseline period, indicating a relatively low signal-to-noise-ratio.

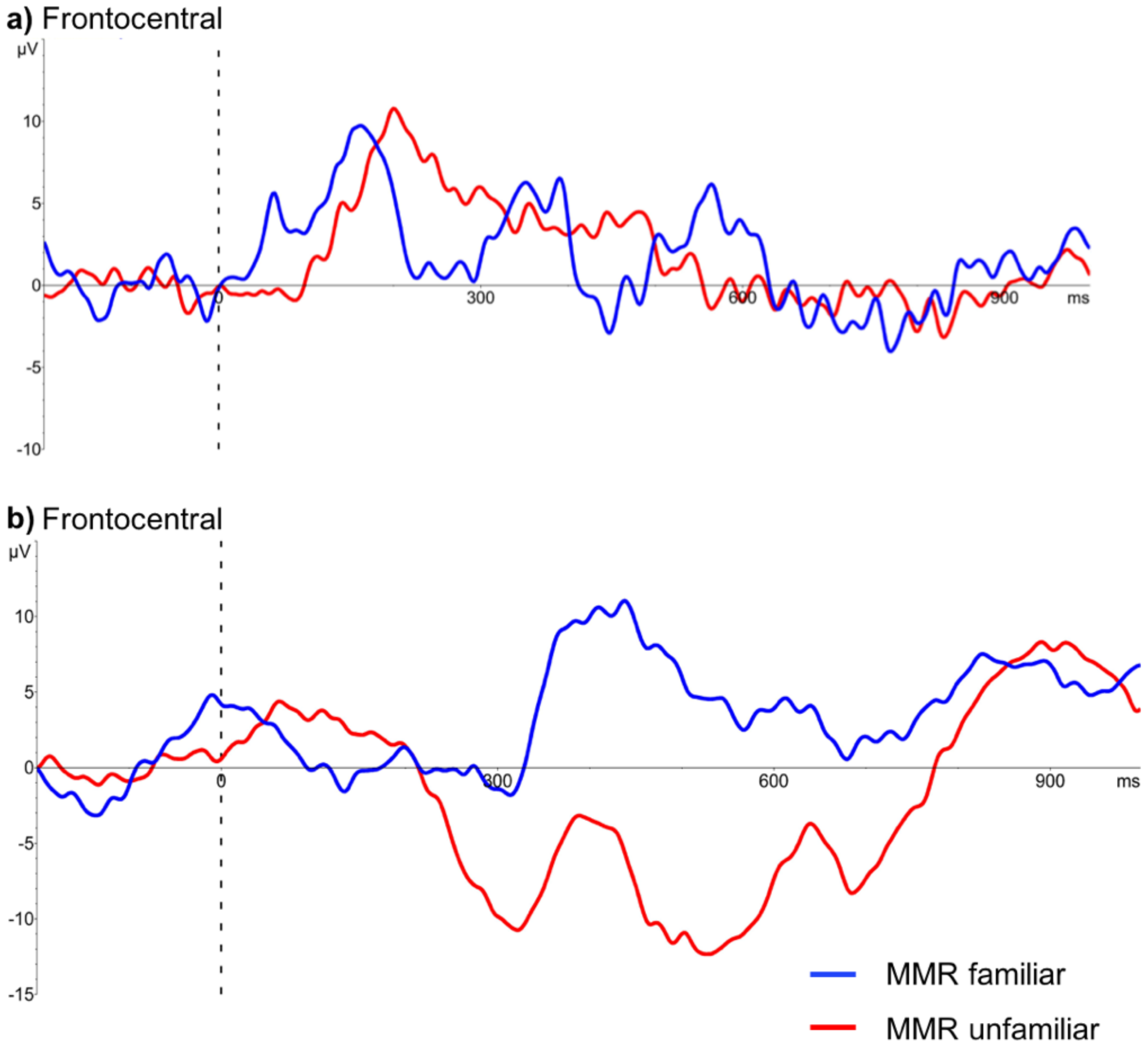

Figure 4.5: Individual mismatch responses.

Representative mismatch responses (MMRs) in the frontocentral region of interest (ROI) from two individual participants, 1 (a) and 2 (b), demonstrating the observed interindividual response variability. MMRs to the familiar voice are colored in red, MMRs to the unfamiliar voice are shown in blue.

4.2.3 Statistical comparison

4.2.3.1 Early measurement window (100-300 ms)

The one-way ANOVA showed that voice familiarity had a significant effect on the area amplitudes of the MMR [$F(1,31)=4.82$, $p= .036$, $\eta_P^2= .14$] measured in the frontocentral ROI in the early measurement window. The means of the area amplitude revealed the direction of this effect: the mean area amplitude of the ERP to the unfamiliar condition (-286 ± 128 μV·ms) was lower than for the familiar condition (180± 150 μV·ms), as shown in Figure 4.6.

4.2.3.2 Late measurement window (300-450 ms)

In the late measurement window, familiarity trended to influence the area amplitudes of the MMRs in the frontocentral ROI [$F(1,31)=3.41$, $p= .075$, $\eta_P^2= .099$], without reaching significance. The area amplitudes of the response to the familiar condition (274 ± 95.8 μV·ms) tended to be greater than to unfamiliar condition (1.30 ± 106 μV·ms), as illustrated in Figure 4.6.

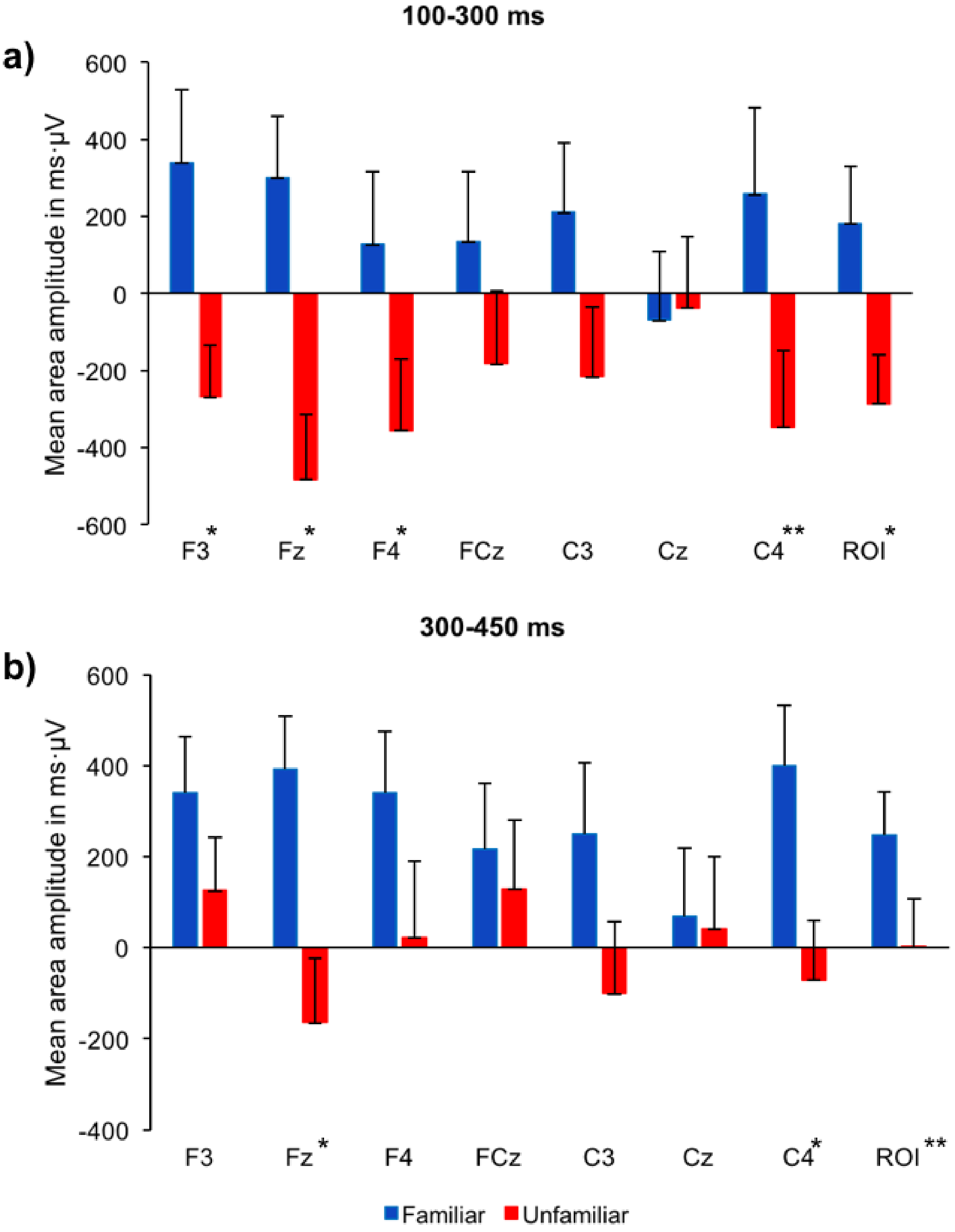

Figure 4.6: Mean area amplitude of the mismatch responses.

Mean area amplitude of the mismatch responses (MMRs) in the measurement windows of 100-300 ms (a) and 300-450 ms (b), demonstrated for each electrode site and for the frontocentral region of interest (ROI). The mean area amplitudes of the MMRs to the familiar deviant (blue) and the unfamiliar deviant (red) condition are contrasted. Error bars represent the standard error of the mean (SEM). Electrode sites that showed a significant difference in area amplitude between the two conditions are marked with * ($p < .05$). Electrodes sites showing an area amplitude that tended to differ are marked with ** ($p < .10$)

4.2.4 Explorative *t*-tests at the single electrode sites

Figure 4.6 demonstrates the mean area amplitudes of the ERPs for the single electrode sites as well as the frontocentral ROI in both measurement windows. Descriptively, the mean area amplitude of the MMR to the familiar deviant condition was more positive than for the unfamiliar deviant condition at all electrode sites in the early measurement window, except at Cz. However, *t*-tests revealed that the differences between both deviant conditions were only significant for F3, Fz and F4 in the early measurement window (*p*< .05).

Descriptively, the mean area amplitude of the MMR to the familiar deviant condition showed to be greater than to the unfamiliar deviant condition at all electrode sites in the late measurement window, as it is depicted in Figure 4.6. However, significant differences were only found for Fz and C4. The results of the explorative t-tests are summed up in Table 4.1.

Table 4.1: Results of the t-tests of the mean area amplitude at the single electrode sites

Results are shown for the early measurement window (100-300 ms) and the late measurement window (300-450). Significant *p*-values < .05 are highlighted with *. Trends (*p*< .10) are labeled with **.

	100-300 ms			**300-450 ms**		
	df	*t*	*p*	*df*	*t*	*p*
F3	29	2.49	.019*	29	1.27	.22
Fz	28	3.08	.005*	28	3.07	.005*
F4	28	2.08	.047*	28	1.46	.16
FCz	25	1.28	.212	25	0.60	.56
C3	27	1.19	.247	27	0.96	.35
Cz	24	0.03	.93	24	0.17	.87
C4	30	1.82	.078**	30	2.45	0.20*

4.3 Analyzing control variables

4.3.1 Testing normality

The area amplitudes of the MMRs to the familiar and unfamiliar deviant condition measured in the frontocentral ROI in the early measurement window reported a normal distribution in the Kolmogorov-Smirnov test. The age in days of our study population also showed to be normally distributed data.

Further, the area amplitudes of the MMRs to the familiar and unfamiliar condition in the late measurement window showed normality. The distribution of sleepiness score inquired before presenting the voice paradigm [*D*(23)= .22, *p*= .004] and the hours of sleep in the last 24 h [*D*(27)= .22, *p*= .002] both violated the assumption of normality. Both were positively skewed in their distributions.

4.3.2 Control variables

No significant relationships between the age in days and the area amplitudes of the MMRs elicited by the familiar as well as by the unfamiliar condition were found for both measurement window.

Spearman's correlation coefficients were not significant for the sleepiness score assessed before presenting the paradigm and the area amplitudes of the MMR to both, the familiar and the unfamiliar deviant condition in the early measurement window. Also, the sleepiness score and the area amplitudes of the MMR to both deviant conditions reported no significant relationship in the late measurement window. The amount of sleep in the 24h before the day of testing was not significantly related to the area amplitudes of the MMRs to both deviant conditions neither in the early nor the late measurement window.

The results indicate that none of the control variables that we inquired did influence the area amplitudes of the MMRs to both deviant conditions in both measurement windows.

5 DISCUSSION

5.1 Main findings

In the present study, we aimed to investigate electrophysiological responses that reflect voice familiarity in infants between two and four months of age. We set the focus on the mismatch response (MMR), an event-related potential (ERP) component that reflects processes of deviance detection and had been found to be modulated by the listener's familiarity with the presented stimulus (Beauchemin et al., 2006; Beauchemin et al., 2011). We recorded ERP responses in infants that were elicited by stimuli arranged in an oddball paradigm. This paradigm consisted of the word 'baby' pronounced by a female unfamiliar voice reflecting a frequently presented standard stimulus, the maternal voice as well as another female stranger's voice, both representing deviant stimuli. The MMR is calculated by subtracting the response to a frequently presented standard stimulus from the response to a deviant stimulus (Cheour, 2007).

In a first step, an analysis of the original ERP responses was performed for exploring the differences between the response to the standard stimulus and the responses to the two deviant stimuli. The analysis of the original ERPs revealed that the response to the mother's voice presented as a deviant stimulus in the oddball paradigm was significantly more positive than the response to the frequently presented standard stimulus in a late measurement window between 300 and 450 ms after stimulus onset. In contrast, the response to the unfamiliar deviant voice showed to be significantly more negative between 100 and 300 ms after stimulus onset but did not show differences in the late measurement window.

For bringing to light differences in the responses that are behind preattentive processes of voice familiarity, MMRs to the familiar deviant (mother's voice) and the unfamiliar deviant (unfamiliar voice) stimuli were computed and analyzed in a second step. The MMR to the mother's voice was significantly larger at the frontocentral electrode sites than the MMR to the unfamiliar deviant voice stimulus between 100 and 300 ms post-stimulus. Between 300 and 450 ms after stimulus onset, the MMR to the mother's voice presented as a familiar deviant stimulus in the oddball paradigm only tended to be more positive.

Altogether, we observed in both, the original ERP responses as well as in the MMRs, differences between the responses to a highly familiar voice stimulus and to an unfamiliar voice stimulus.

5.1.1 ERP Components

For analysis, we chose to measure the area amplitude in two different time windows adhering to the preexisting studies and literature on the MMR in infancy. The early time measurement window, between 100 and 300 ms post-stimulus, aimed to capture the traditional MMN-like response that had often been observed in infants (Cheour, 2007; Kushnerenko, Ceponiene, Balan, Fellman, Huotilaine, et al., 2002), here referred to as early negativity (EN) (termed by Kushnerenko et al., 2013).

We chose the late measurement window, between 300-450 ms post-stimulus, to detect a late positive subcomponent of the MMR that had been detected frequently in infants, here referred to as positive component (PC) (reported by Cheng, Lee, Chen, Wang, & Decety, 2012; Cheour, 2007; He et al., 2009; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002; termed by Kushnerenko et al., 2013).

The analysis of the original ERP responses in the early time window revealed that the responses to the unfamiliar deviant stimulus, represented by a stranger’s voice, was significantly more negative than the responses elicited by the frequently presented stranger’s voice at the anterocentral electrode sites. Surprisingly, the responses to the familiar deviant stimulus, represented by the mother’s voice, did not show a significant difference in the area amplitude compared to the frequently presented unfamiliar voice (standard condition) in the early measurement window.

Contrasting our results, a positive response to the maternal voice than to a stranger’s voice in a similar, but broader time window (150-400 ms) had been found by Mai and colleagues in two-month-old infants. They had observed higher peak amplitudes in response to a mother’s than to a stranger’s voice in the midline electrodes, interpreting it as a correlate for rapid voice discrimination or a process modulated by an attention shift to the mother’s voice (Mai et al., 2012). However, it should be noted that the test condition differed from our study since

Mai et al. used a paradigm in which familiar and unfamiliar voices were presented equally often. Another difference to study is that they used the peak amplitude to quantify the ERP components.

Further, we observed a positive component in response to the mother's voice in the visual inspection of the grand-average original ERP responses (see Figure 5.5.1, column a)). However, in the analysis, the area amplitude in the early time window in response to the mother's voice did not differ significantly from the one in response to stranger's voice that provided the standard stimulus. Nevertheless, it should be considered that two components of diverging polarity might appear in the early measurement window and thus their differing polarities cancel out each other when calculating the area amplitude. This phenomenon might be avoided by choosing narrower time windows when applying the measurement of area amplitude. In a different approach of quantifying the ERP components, we used narrower time windows by dividing the post-stimulus interval between 100 and 400 ms into smaller intervals of 100 ms each (Zinke, Thone, Bolinger, & Born, 2018).

The comparison of the MMRs to the mother's and stranger's voice corroborated the observations made in the analysis of the original ERP responses in the early time window. The MMRs elicited by the mother's voice differed from the MMRs elicited by the deviant stranger's voice; the area amplitude of the MMR to the deviant stranger's voice at the frontocentral ROI was more negative than to the mother's voice. These results are in concordance with the observations made in neonates by Beauchemin and colleagues who found differences in the MMRs to the maternal and a stranger's voice in a latency window of 176-224 ms (2011).

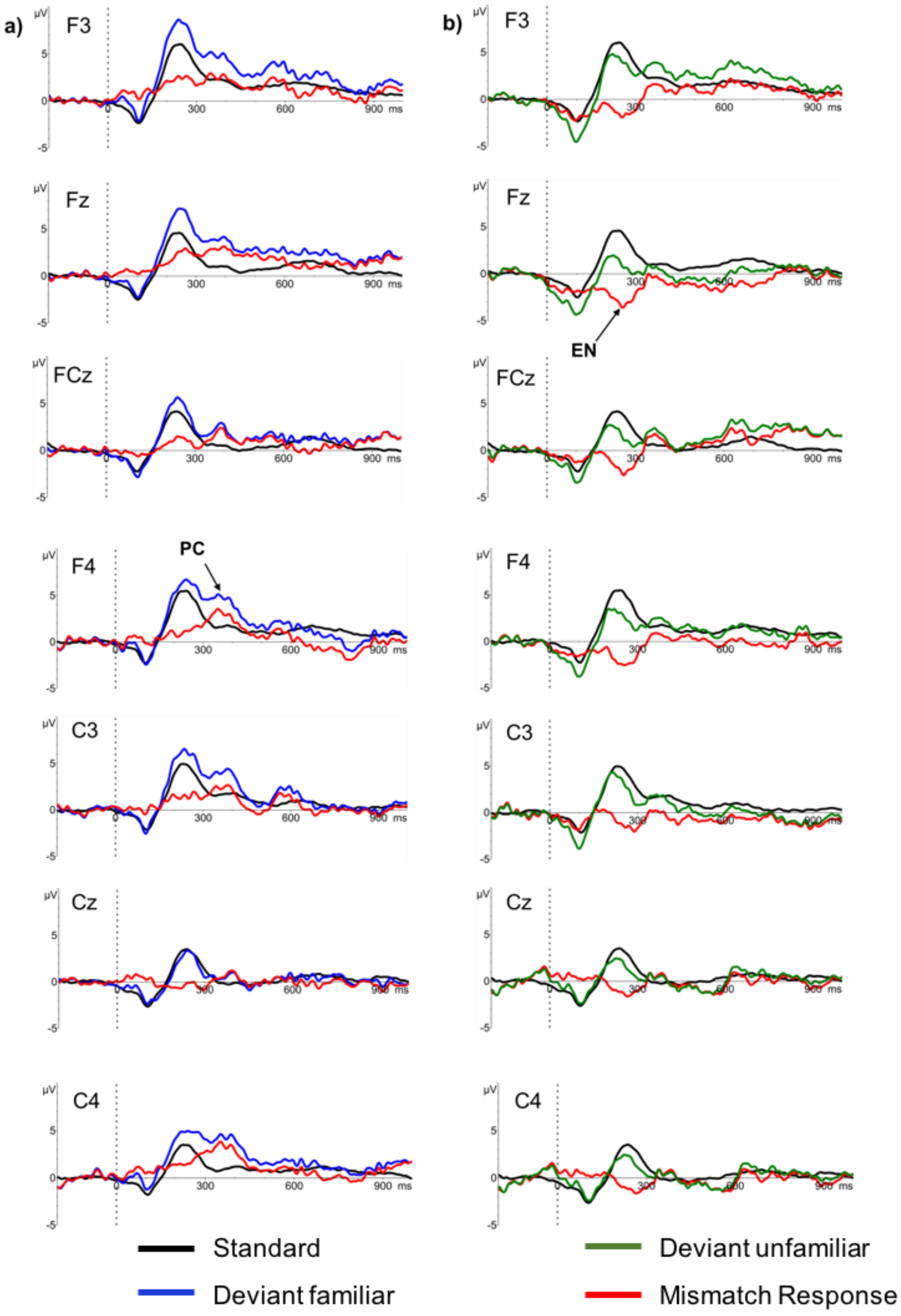

Figure 5.5.1.: Original ERP waves and their mismatch responses.

Grand-average original waves displayed with their mismatch responses (MMR). Column a) shows the original ERP responses to the standard stimulus (black), to the familiar deviant stimulus (mother’s voice, colored in blue) and their calculated MMRs (red). A positive component (PC) in the MMR to the mother’s voice was observed in the late measurement window (300-450 ms). Column b) depicts the original waves in response to the standard (black), to the unfamiliar deviant stimulus (stranger’s voice, colored in green) and their calculated MMR (red). An early negative component (EN) was observed in the MMR to the stranger’s voice.

With a closer look at the MMR grand-averages at the single electrodes depictured in column a) and b) of Figure 5.5.1 and Figure 4.3, wave deflections come to light that might have led to the described differences in area amplitude in the early time window. At all electrode sites, a negative deflection (peaking at around 250 ms) in the MMRs to the unfamiliar deviant voice stimulus was found (see column b). This negative deflection was less distinctive at the central than the electrodes. In all electrodes, except Cz and C3, an additional earlier negative deflection (around 100 ms) was observed in the MMR to the unfamiliar deviant condition in the early time window. A double negative peaking in the early time window was also distinguishable when averaging the frontocentral electrodes to a frontocentral Region of interest (ROI) (Figure 4.4). However, none of these negative peaks was observed in the grand-average MMR elicited by the mother's voice.

Taking the negativity of the area amplitudes of the MMRs and the visually observed deflections in the early time window into consideration, these results might be interpreted as the elicitation of an early negativity (EN) in response to the unfamiliar deviant stimulus. Beauchemin and colleagues observed in neonates that the responses to an unfamiliar deviant voice and a frequent unfamiliar voice tended to differ in the midline electrodes around 176 ms after stimulus onset, however, this difference did not reach significance (Beauchemin et al., 2011). The distinguishable EN in response to the unfamiliar voice observed in our study leads to the conclusion that this subcomponent of the MMR matures in the first three months of life, probably after the infant has gained a certain amount of voice experience.

As described in the introduction, the EN was reported to be elicited by acoustic deviance within a stream of frequently presented stimuli between the age of two and four months and had been interpreted as a precursor of the adult mismatch negativity (MMN) (He et al., 2009). The EN in the MMR to the unfamiliar deviant stimulus that we observed possibly reflects a process of preattentive auditory deviance detection. It can be considered as a 'comparison process' of the short-term memory system between the neural trace formed by the unfamiliar deviant stimulus and the memory trace that had been formed by the preceding presentation of the standard stimulus (Cheour et al., 2000; Näätänen, 2012; Zinke et al., 2018).

The original ERP responses to the mother's voice reported significantly more positive than the responses to the stranger's voice used for the standard condition in the late measurement window. A later positive component following the P2 component in response to the mother's voice might have contributed to these results. This later positive component in response to the familiar deviant stimulus was also visually distinguishable in the grand-average MMRs at the single electrode sites and in the frontocentral ROI. However, the area amplitude of the MMRs to the familiar deviant stimulus only tended to be more positive than the area amplitude of the MMRs to the stranger's voice in the analysis.

However, when comparing the mean amplitude of the MMR in a narrower time window between 300 and 400 ms post-stimulus in a different approach of ERP quantification, the MMR to the familiar deviant stimulus revealed to be more positive than to the unfamiliar deviant stimulus. We found a robust positive deflection in response to the mother's voice in this time window (Zinke et al., 2018).

We cautiously interpreted that this trend in the comparison of the MMRs and the significant difference observed the original ERP responses between 300 and 450 ms might reflect the elicitation of a positive component (PC) which had been described as a part of the infant MMR before (see Mismatch response (MMR) in infancy). In contrast, no significant difference in area amplitude was observed between the original ERP responses to the unfamiliar deviant voice and the standard stimuli, suggesting an absence of a PC to the unfamiliar voice in the late measurement window.

There are different explanatory approaches for the absence of an EN component in the MMR to the mother's voice when presented as deviant stimuli in an oddball sequence. Firstly, the maternal voice did not elicit an EN at all. Secondly and more likely, the mother's voice did elicit an EN, but it might have been masked by an overlapping component that occurred at a similar or later latency span. A frequently observed positive deflection following the EN in latency is the positive component (PC) that is often reported as part of the infant auditory deviance detection response (Cheour, 2007; He et al., 2009; Trainor, 2008). The phenomena of the PC masking the EN had been described previously by some authors that studied MMRs elicited by pitch changes of the deviant stimulus in longitudinal study designs. According to them, overlapping especially seems to occur at an age between three and six months, an age span that coincides with the targeted age group of our study (Jing & Benasich, 2006; Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002). However, the phenomenon of overlapping components might clarify the absence of EN in the MMRs of the mother's voice, but it does not serve as an explanation why a similar PC was not detectable in response to the unfamiliar voice.

When looking at studies that explored the influence of voice familiarity on the auditory deviance response, an explanatory approach might be found in the interpretation of the observations made by Beauchemin and colleagues. They observed a significant difference in amplitudes of the MMN and the P3a component between a familiar and the unfamiliar voice condition in adults. Especially the P3a component showed a higher amplitude in response to a familiar voice stimulus and thus might haven be modulated by the listener's familiarity with this voice stimulus (Beauchemin et al., 2006).

In a follow-up study in newborns, Beauchemin et al. found that the MMRs to the deviant mother's voice were more positive than to the deviant stranger's (see The influence of voice familiarity on ERP components in early childhood). Beauchemin and colleagues proposed that a P3a-like component might lie behind the observed positivity in response to the mother's voice reflecting a switch of attention to the deviant stimulus. They further proposed that its amplitude was modulated by the familiarity with the mother's voice resulting in the phenomenon of overlapping described before (Beauchemin et al., 2011). This suggestion also can be applied to the observations made in our study: the PC might reflect P3a-like component that is elicited only or higher in amplitude by the mother's voice.

Hypothesizing that the EN and the PC were electrophysiological correlates of different neural processes in infancy, the PC has also been interpreted as a precursor of the P3a component by other authors (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002; Kushnerenko et al., 2013). The P3a component has been observed in adults (e.g., Courchesne, Hillyard, & Galambos, 1975;

Frank, Yee, & Polich, 2012) and in children (Ceponiene et al., 2004) in response to a deviant stimulus in a sequence of frequent stimuli.

The processes behind the P3a component are usually reported to be different to the processes involved in the elicitation of an MMN. However, P3a is similarly observed in response to a deviant stimulus within an oddball sequence. Different theories on the neural processes that are related to the unintentional novelty detecting response of the P3a have been formulated: 1) Donchin and Coles suggested that the P3 component reflects to a process of contextual updating (1988), 2) others suggested the P3a being part of the orienting response or/ and 3) attentional capture (Friedman, Cycowicz, & Gaeta, 2001).

An orienting response toward the mother's voice, even prenatally, has been observed in several behavioral studies, as it is described in the introduction (see Role of the mother's voice in infancy). However, a recent study by Moon and colleagues showed diverging results in neonates. Through a behavioral study, they aimed to contemplate the results of electrophysiological studies on voice familiarity by using the same type of voice stimuli that had been applied previously to elicit ERPs in neonates and that was also used in our study. They compared the sucking behavior of neonates in response to the word 'baby' pronounced by the mother's voice and by a female stranger's voice. Surprisingly, their results did not demonstrate a difference in responses to the two stimuli. They interpreted the absence of an overt behavior as a lack motivation to react to the maternal voice and not as a contradiction to a voice recognition process to the maternal voice for which electrophysiological studies had provided evidence previously (Moon et al., 2015).

Another possible positive-peaking component that might have overlapped the EN, with a similar electrode distribution and temporal resolution as the PC, could be a precursor of Fronto-Temporal Positivity to Voices (FTPV) component that had been observed adults and preschool children (see Voice recognition in adults). Thus, the maternal voice might be processed in voice-specific pathways, whereas the response to a stranger's voice might be processed not voice-specifically but as any other acoustic stimulation. In line with the hypothesis of the presence of a voice-specific response to familiar voices in early childhood, Beauchemin and colleagues found that the maternal voice activated predominantly voice-specific areas in neonates (2011). However, there does not exist any MMR study up to date that examines the responses to human-made sounds in comparison to non-human sounds in infants.

As outlined before, there exist only speculations about the possible neural processes behind the PC in infants. The observed positivity elicited by the maternal voice in both measurement windows in our study could have been caused by an attentional switch and/or an orientating response towards the maternal voice, as it had been suggested by Beauchemin and colleagues (2011).

Irrespectively of its possible adult counterpart, the observation of a PC in the MMR to a highly familiar voice but not in the MMR to an unfamiliar voice fits nicely the suggestion made by other authors, that long-term memory templates to the mother’s voice exist in early infancy (deRegnier et al., 2002; Purhonen et al., 2005). These long-term representations, as a part of auditory recognition memory, seem to modulate processes that occur at a preattentive, sensory level (Beauchemin et al., 2011). An influence of these long-term memory representations on the formation of sensory memory traces had been deduced in ERP studies that explored the influence of familiarity on the MMR (for a detailed description see MMN as a tool for detecting voice familiarity). Thus, the PC that we found as a part of the MMR to the mother’s voice might reflect a kind of retrieval of the newly formed sensory memory traces on long-term memory representation for the maternal voice. This mechanism might allow the infant to evaluate if a stimulus is of relevance rapidly after exposure to a stimulus (Zinke et al., 2018). As described in the introduction, the mother’s voice represents such a stimulus of high relevance to the infant (see Role of the mother’s voice in infancy).

To get a better understanding of how the neural processes of voice familiarity develop, it could be of interest to explore if familiarization of an unfamiliar voice changes the MMR elicited by it, or even approximates the MMR elicited by the mother’s voice. Following this idea, Partanen and colleagues provided evidence that familiarization of speech sounds influences their subsequent discrimination, measured as MMR, even when these sounds had been learned prenatally (2013).

To sum up, our results show that the polarity and the subcomponents of the MMR are modulated by the familiarity with the presented deviant stimulus at the age between two and four month. An unfamiliar deviant stimulus, here reflected as a stranger's voice, evoked an EN, reflecting preattentive auditory deviance detection (Cheour, 2007). Contrasting, a highly familiar stimulus, such as the mother's voice, likely elicits a later PC when presented as a deviant stimulus in an oddball sequence, suggesting a modulation of sensory memory processes by long-term memory representations of the familiar stimulus (Beauchemin et al., 2011; Zinke et al., 2018).

Our results are in line with the observations of the only study thus far made on the influence of stimulus familiarity on the MMR in neonates that the mother's voice produces a more positive MMR (Beauchemin et al., 2011). In contrast to the newborns in Beauchemin's study, our infants showed a robust EN in the post-stimulus interval of 100-300 ms, corroborating that this subcomponent consolidates in the first month of life (Jing & Benasich, 2006; Kushnerenko et al., 2013).

5.1.2 Response variability

In our sample of subjects, a broad ERP response variability was observed, a phenomenon that often occurs in infant ERP studies (Coch & Gullick, 2012), thus our results should be interpreted with caution. Reasons for the described variability lie in between-subject and within-subject differences in ERP responses.

De Haan suggested that an important source of differences in ERP responses between subjects of a sample is age. It had been observed that consistency of brain responses increases in the course of development (de Haan, 2007). In fact, it was reported that not all infants show an MMN-like EN. Kushnerenko and colleagues found an EN in 75% of their participants between birth and one year of age (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002), while other authors reported evidence for an EN in only 50%-75% of the newborns (Cheour, 2007). As described in the introduction, age exerts an influence on the MMR and therefore might act as a possible confounder in the present study.

The age range is recommended to be kept as small as possible when researching with ERPs in early infancy. Picton et al. suggested that the age range should be kept between one and two months in infant populations (2000). DeBoer set a narrower age range of ten days as a criterion for infant ERP studies in a more recent publication (2007). With a difference of fifty days of age between the youngest and oldest participant, the age range of our study sample fulfilled Picton's age guidelines for developing study populations. A statistically significant relationship between age and area amplitude for the used measurement windows was not found in our study. Therefore, age as a contributor to the between-subject variance of the area amplitude is found to be unlikely.

Beside the described sources, “day-to-day variability” (p. 478) has been described to contribute to heterogeneous ERP responses within and between subjects of developmental populations by Coch and Gulick (2012).

Another source of variable responses between subjects and within subjects is suggested to lie in the relatively high rate of trial rejection that researchers usually must cope with when assessing ERPs in infants (Coch & Gullick, 2012; Picton et al., 2000). Since a differing number of trials between two tested conditions might lead to a less reliable analysis (Coch & Gullick, 2012), this point also affects the within-subject design used in our study. However, the number of epochs for both deviant conditions (‘familiar vs. unfamiliar’) did not demonstrate a significant difference in our study. The problem of differing numbers of epochs gained significance in the present study when comparing the deviant conditions to the standard condition, that occurred per se with a higher probability in the oddball paradigm. Luck stated that a differing number of trials between conditions does not affect the statistical reliability when an unbiased measure (e.g., mean amplitude) is applied to the waveforms (2014a). In our study, area amplitude was used as a method that follows Luck's principle of unbiased measure.

Some authors have used the number of trials per subject as a covariate in the statistical analysis based on the assumption that a more immature level of neurocognitive development might decrease the number of trials. Excluding these infants from the analysis might lead to biased results by including only subjects that exceeded the required number of trials in the analysis (Coch & Gullick, 2012; Stahl, Parise, Hoehl, & Striano, 2010). Researchers have tried to find ways of trial rejection that detect and reject artifacted trials but at the same time exclude fewer participants as possible to gain data of a representative sample for the targeted age group (Fujioka et al., 2011; Stahl et al., 2010). Independent Channel Rejection (ICR), that was used as artifact rejection method in our study, might provide an approach to this problem and is discussed as a method (for further discussion see Independent Channel Rejection).

Another source of confounds might lie in the state of arousal of the participants during the testing session. In the present study, the ERPs were measured in a state of quiet alertness and a simple sleepiness score was utilized to assess the state of the infant before displaying the oddball paradigm. No relationships between the sleepiness and the area amplitude of the MMRs were found. However, this method might be prone to subjectivities of the mother and to differences in the mother-child relationship. For a follow-up study, we thus suggest recording the infants' EEG rhythms to determine their state of alertness with an objective and more reliable method.

Moreover, Luck suggested that the amount of sleep acquired by the subject before recording ERPs might influence their waveforms (2014b, pp. 249-282). Testing this assumption, no correlations between the hours of sleep in the 24 h before recording the ERPs and the area amplitude of the MMRs were observed.

A further reason for the inter-individual variability might have been caused by clinical differences between the infants. Electrophysiological studies using the maternal voice to assess auditory recognition memory showed that ERP components can differ in impaired study populations. Prematurity (Therien, Worwa, Mattia, & deRegnier, 2004) and an perinatal iron deficiency in infants of diabetic mother's (deRegnier et al., 2007; deRegnier et al., 2000; Siddappa et al., 2004) were reported to influence ERP components that are involved in recognition of the mother's voice. In fact, four mothers of our sample reported a dietetically adjusted gestational diabetes. It is unclear to which extend these perinatal impairments influenced the MMR to familiar voices. Most of the pediatric clinical studies involving the MMR targeted older children with defined clinical manifestations (e.g., autism) since infant populations do not yet show specific symptoms (Cheour, 2007).

An example how predisposition for impairment can influence the MMR in early infancy was given by van Leeuwen and colleagues. MMRs to speech syllables that are difficult to distinguish by persons with dyslexia was reported to be absent in two-month-old with a family risk of dyslexia. Their results indicated that a discrimination response to these syllables are already absent in early childhood (van Leeuwen et al., 2006).

5.2 Discussing methods

ERP research in infants is still at an early stage and many studies differ substantially in terms of methodology. During the testing sessions, we made the experience that recording EEG with infants requires certain spontaneity and that many decisions on methodology were taken based on problems that we encountered while conducting the study as well as during the process of analyzing our results. To gain a better understanding of these problems and for finding solutions to them, a discussion on crucial points of the methodology used in the present study follows.

5.2.1 Voice paradigm

Our findings mainly corroborate the results found in neonates by Beauchemin and colleagues (2011) and extend the knowledge of how voice familiarity influences the MMR, a tool to measure acoustic deviance, in the lifespan between two and four months. However, it should be noted that there exist differences between the study conducted by Beauchemin and the present study besides the age difference between the study populations. In their study, the ERPs were recorded while the neonates slept, contrasting our study design in which ERPs were recorded during quiet alertness. Differences in the state of alertness might lead to varying study results on the MMR in infancy (see Mismatch response (MMR) in infancy).

Furthermore, the unfamiliar voice in Beauchemin's study was provided by a nurse that had accompanied the mother throughout pregnancy and thus was not completely unfamiliar to the newborn. Beauchemin and colleagues chose to use a not completely unfamiliar voice as deviant stimulus to avoid contamination of processes that involve novelty detection. Contrasting, we selected a highly unfamiliar voice as the unfamiliar deviant stimulus. Another difference to the study conducted by Beauchemin lies in the stimuli used in the oddball paradigm. Beauchemin and colleagues used the single vowel ‘a’ (2011), whereas the stimulus in our oddball paradigm contained two syllables and thus was of a more complex acoustic nature. A word with two syllables might possibly provide a more suitable stimulus when testing voice recognition since it contains richer voice features.

A possible confounder in the paradigm could have been crossmodal integration since the participants saw the mother’s face during the presentation of the paradigm, but not the corresponding faces to the other stimuli types. However, the presence of the mother while presenting the paradigm was a crucial attention attractor keeping the infant calm. The maternal face might have elicited a brain response itself that could have interfered with the ERP signal (Hoehl & Wahl, 2012). However, we tried to keep sources of crossmodal integration as small as possible, since the mother did not move their mouths or spoke during the presentation. Additionally, emotional integration as possible confounder was avoided when recording the voice stimuli: the voice stimuli were recorded without emotional connotation and the speakers were instructed to use adult-directed speech. In contrast, infant-directed speech, so-called ‘motherese’ differs in many acoustic aspects from adult-directed speech. It is slower, of a positive prosody and higher in pitch (Fernald, 1985).

In adults, the MMN is reported to be prone to be contaminated by differences between the obligatory auditory ERP components elicited by the standard and the deviant stimuli and “those in the refractoriness of the neural populations activated by the two stimuli because of the probability difference” (Näätänen, 2012, p. 144). The effect of refractoriness is assumed to be provoked by an imbalance of stimulus probability in the composition of an oddball paradigm since the standard stimulus is usually presented with a higher probability than the deviant stimulus. It is assumed that due to refractoriness of frequency-specific neurons that are activated, a stimulus presented as a standard stimulus elicits a response lower in amplitude than if it would be presented as a deviant stimulus, with a lower probability of occurrence (Kujala, Tervaniemi, & Schroger, 2007).

A salient cue about the role of refractoriness of the infant MMR was provided by Háden and colleagues. Their research team investigated the influence of refractoriness on the MMR in neonates. They found differences in responses to a tone presented as deviant in an oddball paradigm with a low probability of presentation and the same tone presented in a sequence of several differing tones. Thus, the response to a deviant stimulus in an oddball paradigm cannot be fully explained by “frequency-specific refractoriness the neural generators” in newborns (Háden et al., 2016, p. 26). In their longitudinal study on the auditory deviance detection in infancy, Kushnerenko et al. integrated a similar control paradigm besides the usual oddball paradigm. The control paradigm was suggested to enable differentiation between the influence of the genuine deviance detection and the effect of refractoriness on the ERP response. In the control paradigm, the stimulus representing the deviant in the oddball paradigm was embedded between other tones, all occurring equiprobably. To assess the contribution of the genuine auditory deviance detection to the response, the responses of the deviant stimulus was then compared between the two paradigms (Kushnerenko, Ceponiene, Balan, Fellman, & Näätänen, 2002). The application of a similar control paradigm could be reasonable in a follow-up study on the MMR to voice familiarity, since voice-tuned regions might be affected by a similar effect of refractoriness.

5.2.2 Independent Channel Rejection

As mentioned above, researchers conducting infant ERP studies are often confronted with a high drop-out rate of trials since infant recordings show a higher rate of artifacts and the recording time should be kept as short as possible due to limited cooperativeness of the young participants (Fujioka et al., 2011; Hoehl & Wahl, 2012; Picton et al., 2000). A rejection rate of 40% of the trials or more is likely to be seen in infants (Picton et al., 2000).

Particular in view of the diverse results reported on the MMR in infancy (see Introduction), Trainor suggested that “the results obtained will depend greatly on the criteria by which infants are included and excluded into the analysis” (Trainor, 2008, p. 90).

In our study, we also faced the problem of a high drop-out rate: artifact rejection and the criterion of a minimal number of ten epochs for the deviant conditions resulted in a rejection rate of 44.57% of the obtained epochs. Probably the rejection rate would have been even higher when rejecting whole trials that showed artifacts in at least one channel. Since we used an oddball paradigm, the number of epochs for the deviant conditions was more limited than it would be for another type of paradigm, aggravating the problem of the lack of trials to the deviant conditions provided for averaging (Fujioka et al., 2011). Additionally, recording ERPs in a state of wakefulness, as it was done in the present study, enhances the number of artifacts in a study population that is already prone to produce more artifacts.

To increase the number of included subjects as well as the number of included epochs per condition and along with that the statistical power of our study, we applied ICR on our data (Fujioka et al., 2011; He, Hotson, & Trainor, 2007; He et al., 2009). The standard approach in adult and in most of the infant ERP studies, Conventional Trial Rejection (CTR), involves elimination of all channels of a trial that might show an artifact in only one or more of the channels. However, when applying ICR, epochs of artifact-free channels are included in the process of averaging also when they were recorded in a trial that included other artifact-contaminated channels.

Fujiko et al. developed ICR as a method of artifact rejection for infant ERP studies to optimize the number of epochs for averaging. They stated that the CTR used in adult ERP research is not an appropriate approach when preprocessing infant data. Spontaneous effusive movements often occur in infant EEG recordings that cause reduced skin-to-electrode contact resulting in temporary artifacts in single electrodes, but do not affect all channels (Fujioka et al., 2011). Trainor stated that “the sources of the artifact are likely somewhat different and more independent in infants than in adults” (Trainor, 2007, p. 82), corroborating the assumption that ICR might be favorable when processing artifacts in infant ERP data. Appling CTR on data with a high rate of this type of artifacts might not represent sufficiently the ERP signal. Supporting these assumptions, Fujioka and colleagues found in their study that the ICR procedure is more effective in the estimation of the original ERP signal than the CTR method when performing artifact rejection on infant data (Fujioka et al., 2011).

However, Luck criticized that the electrode-specific rejection distorted the experimental effect since segments remained in the analysis that still contain low artifacts. He stated that the electrode-specific approach led to a differing number of trials between the condition that remain for the process of averaging and that it complicated the determination of scalp-distributions (Luck, 2014b, Chap. 6). In our study, we bypassed the problem of a different number of trials per by using an unbiased measure (see above). Nevertheless, the problem of unprecise spatial scalp-distribution does not apply to our analysis of the MMR ran on the frontocentrally clustered averages since it did not target information of spatial distribution.

CTR was applied to our dataset as a possible approach for the analysis. The CTR approach, however, resulted in fewer participants (n=22) and a reduced mean number of trials per condition (*M*(familiar): 23.38, *SD*: 6.17, *M*(unfamiliar): 23.57, *SD*:6.68) remaining for averaging. Furthermore, this approach led to a visible increase in the SEM and noise of the grand-average MMRs. Thus, we evaluated the CTR to be a better fit for our dataset enabling a more representative sample by including artifact-free epochs from more participants.

5.2.3 Choosing the area amplitude for quantifying the MMR

The great variety of results labeled as MMR and the diverging criteria set for its detection in studies on the MMR in infancy can be challenging when analyzing and interpreting one’s own results. The great variability of the reported study results on one hand and the interindividual differences one faces studying a developing population, on the other hand, complicate the acquisition of a priori information on a suitable measurement window or quantifying method when researching with the infant MMR.

Two examples illustrate the differing electrophysiological definitions used for quantifying the MMR: Leppanen and colleagues quantified the MMR in newborns as “the amplitude of the largest deflection (regardless of the sign) of the difference wave within a time window of 150-375 ms in relation to the 50 ms prestimulus baseline” (2004, p. S94). Contrastingly, Kushnerenko and colleagues defined the

MMR in infants between neonatal period and the age of twelve months as “the largest negative deflection in the difference waveform between 80-300 ms after stimulus onset, greater than the average baseline voltage by 1.0 µV at any of the four fronto-central electrodes” (2002, p. 1844).

However, bracing criteria for quantifying the targeted ERP component on previous studies is the best way to gain reliable results (Luck, 2014b, Chap. 10). In the present study, we decided to use area amplitude to quantify the MMR, meaning that areas under the baseline were subtracted from areas above the baseline, thus the resulting area can have negative values. The area amplitude results when multiplying mean amplitude with the duration of the chosen measurement window (Luck, 2014b, Chap. 9). No study to date has ever used amplitude area measures to quantify MMR in infancy. The advantages of the area amplitude, are that it neither does defines polarity, nor does it depend on peak criteria. Thus, it might represent a promising tool to measure an ERP component whose polarity and wave configuration has shown manifold facets in past studies, such as the MMR (see Mismatch response (MMR) in infancy). A disadvantage might lie in the fact it is prone to the problem of overlapping components, as is the mean amplitude (Luck, 2014b, Chap. 9).

A limitation of the present study is, that no latency measures were applied. No differences in latency between the MMRs to the mother’s and the stranger’s voice were not explored in our study. For instance, Purhonen and colleagues found that a P150 and a N250 component as part of auditory ERP components were shorter in latency to the mother’s than to a stranger’s voice (2005). Thus, voice familiarity might also have an impact on the latency of the MMR. Within the scope of the area measures, ‘fractional area latency’ for estimating the midpoint latency (2014b, pp. 296-299, Chap. 9) could be to an appropriate measure to investigate latency differences (Luck, 2014b).

5.2.4 Electrode sites

For the analysis of the MMRs, the frontocentral electrode sites were averaged together resulting in a frontocentral Region of interest (ROI). This averaged channel was used to assess the area amplitude of the MMRs to the deviant stimuli. The ROI was adapted to studies that investigated the MMR in infants and specifications in the literature (Alho et al., 1990; Cheour, 2007; Mai et al., 2012; Partanen et al., 2013).

However, specifications about the scalp location of the MMR in infancy diverge: Beauchemin and colleagues used an averaged frontocentral midline cluster to measure the MMR (Fz, FCz, Cz) (2011). In our result, the explorative t-test and visual inspection of the MMR waves revealed differences between the MMRs to the unfamiliar deviant and the mother's voice for both left and right hemisphere electrodes. Excluding these electrode sites from the analysis would not reflect a representative signal of the MMR. Furthermore, the approach of averaging across an electrode cluster has advantages over a single electrode analysis: using electrode site as an ANOVA factor, produces a higher amount of *p*-values that results in an increased familywise error rate and consequently inflates the probability of Type 1 error.

Furthermore, averaging across an ROI in which the main effect is expected reduces the level of noise in the waveforms. The disadvantage of this approach is a loss of information of the scalp distribution of the main effect (Luck, 2014b, Chap. 10). In the present study, the process of averaging across the frontocentral electrodes created a broad electrode cluster ruling out any statements on the scalp distribution of the found MMR components. In the analysis of the original ERP responses, the scalp site was included as a variable, however, no interaction with the condition ('standard vs. deviant') was found, indicating that the responses to the stimuli were not influenced by the scalp site at which it had been measured. Our findings of a broad frontocentral scalp distribution of the MMRs corroborates the observation made by Cheour. This phenomenon is possibly traced back to a smaller anatomical distance of brain regions and lower bone thickness in infants leading to conductance differences (Cheour, 2007).

A follow-up study with a high-density system, e.g., geodesic sensor nets (GSN), and a higher number of trials might shed light on the scalp distribution of the found MMR components. Moreover, GSNs bear no risk for salt bridges caused normally by conductive gel that it is required when using an EEG-cap system (Hoehl & Wahl, 2012; Johnson et al., 2001).

We further recommend considering the position of the infant while recording the EEG signal, since the supine position likely causes artifact in the parietal and occipital electrodes as it was observed for the Pz electrode in the present study. In other studies, on auditory ERPs, the infants were seated on their mothers’ laps while listening to the paradigm (Hoehl & Wahl, 2012; Mai et al., 2012) or held by them (Purhonen et al., 2005).

6 ABSTRACT

Our study investigated the influence of voice familiarity on the mismatch response (MMR) in infants between the age of two and four months (*M*: 98.53 days). The Infant MMR is an event-related potential (ERP) component that reflects processes of preattentive deviance detection within a stream of frequent stimuli and is often described as a precursor of the adult mismatch negativity (MMN). In the past, the infant MMR and the adult MMN have been used to detect neural correlates for processes that lie beyond voice recognition.

We targeted to elicited ERPs and MMRs by an oddball paradigm that was presented to the infants. It consisted of the word ‘baby’ pronounced by three different female speakers (stimulus duration 400 ms, interstimulus interval 600 ms, total duration 10 min). The stimuli were presented as 1) a ‘standard stimulus’ (frequently presented stranger’s voice), 2) a ‘familiar deviant stimulus’ (infrequently presented mother’s voice) and 3) an ‘unfamiliar deviant stimulus’ (infrequently presented stranger’s voice).

For quantifying different subcomponents of the MMR and ERPs, the area amplitudes were determined for an early measurement window (100-300 ms post-stimulus) targeting an MMN-like early negativity (EN) and for a late measurement window (300-450 ms post-stimulus) capturing a positive component (PC).

As a pre-analysis, the original ERP responses to mother’s and stranger’s voice were compared to the responses to standard stimulus at the frontocentral electrode sites in the two targeted latency windows. The area amplitude of the original ERP responses to the stranger’s voice showed to be more negative than the area amplitude of the responses to the standard stimulus in the early time window, indicating the elicitation of an EN component. However, the area amplitude of the original ERP responses to the mother’s voice was more positive than in response to the standard stimulus between 300 and 450 ms, indicating the presence of a PC.

In the second step, differences between the MMRs to the mother’s and the stranger’s voice were explored. Between 100 and 300 ms, the area amplitude of the MMRs to the mother’s voice was significantly larger than the area amplitude of the MMRs to the stranger's voice at a frontocentrally averaged scalp site. In the late measurement window, the area amplitude of the MMR to the mother's voice tended to be more positive than to the stranger's without reaching significance. Our results corroborate the findings in neonates of Beauchemin and colleagues that a highly familiar voice stimulus produces a positive MMR (Beauchemin et al., 2011).

Our findings suggest that an early MMN-like negative component (EN) was elicited in response to an infrequently presented unfamiliar voice. We interpreted this early component as an auditory deviance detection response that occurs preattentively. Since a EN had not been found in neonates (Beauchemin et al., 2011) in response to a deviant voice stimulus, our findings showed that a stable auditory deviance detection response to voice stimuli likely emerges in the first months of life, respectively, when the infants have already gained a certain amount of voice experience. The absence of a similar component in response to the infrequently presented mother's voice might be explained by the presence of the PC found in the late measurement window that might have overlapped the preceding EN. We carefully interpreted our findings in a way that this positive deflection might be elicited due to a shift of attention towards the highly familiar voice. Thus, the PC might represent a precursor of the adult P3a component. An attentional shift towards the mother's voice could be explained by the fact that it represents a highly relevant stimulus in early infancy.

Further, we interpreted the presence of a PC in response to a highly familiar voice as a reflection of processes of auditory recognition memory modulating processes that are part of the preattentive auditory deviance detection. From our observations, it can be deduced that infants at the age of two-to-four months process a familiar voice differently than an unfamiliar one. A part of the processes behind voice recognition in early childhood presumably occurs preattentively and they seemed to be altered to some extend by long-term memory representations of the auditory recognition system (Beauchemin et al., 2006; Beauchemin et al., 2011).

7 ZUSAMMENFASSUNG

Die vorliegende Studie untersuchte den Einfluss von Stimmenbekanntheit auf die Mismatch Response (MMR) bei zwei bis vier Monate alten Säuglingen (*M*: 98,5 Tage). Die im Neugeborenen- und Säuglingsalter beobachtete MMR ist eine Komponente, die zu den Ereignis-korrelierten Potentialen (EKPs) zählt und oft als Vorläufer der im Erwachsenenalter vorkommenden Mismatch Negativity (MMN) beschrieben wird. Der MMR liegen präattentive Prozesse zugrunde, Abweichungen in einer Folge von regelmäßigen Stimuli zu detektieren.

Unsere Studie hatte zum Ziel EKPs und MMRs mittels eines Oddball-Paradigmas bei unseren Probanden hervorzurufen. Dieses Paradigma bestand aus dem Wort “Baby“, das von drei verschiedenen weiblichen Stimmen gesprochen wurde (Stimulusdauer 400 ms, Interstimulusintervall 600 ms, Gesamtdauer 10 min). Die Stimuli wurde den Probanden als 1) ein Standardstimulus (fremde, sich wiederholende Stimme), 2) ein bekannter, abweichender Stimulus (selten vorkommende mütterliche Stimme) und 3) ein unbekannter, abweichender Stimulus (fremde, selten vorkommende Stimme) dargeboten. Die Messung erfasste zwei Zeitfenster: Das frühe Zeitfenster (100 und 300 ms nach Stimulusbeginn) hatte zum Ziel, eine frühe negative Komponente (FN) zu erfassen. Das spätere Zeitfenster (300 und 450 ms nach Stimulusbeginn) diente dazu eine späte positive Komponente (PK) zu untersuchen.

In einem ersten Analyseschritt wurden die ursprünglichen EKPs in Reaktion auf die bekannte abweichende Stimme und auf die fremde abweichende Stimme des Standardstimulus miteinander verglichen. Für das frühe Zeitfenster zeigte sich, dass die FA der EKPs als Antwort auf die fremde abweichende Stimme in den frontocentralen Elektroden negativer ausfiel als die FA der EKPs als Antwort auf die Stimme des Standardstimulus. Die EKPs als Antwort auf die mütterliche Stimme zeigten Unterschiede zu den EKPs des Standardstimulus im späten Zeitfenster.

In einem zweiten Analyseschritt wurden die MMRs der bekannten abweichenden und der fremden abweichenden Stimme miteinander verglichen. Im Zeitfenster von 100-300 ms fiel die FA der MMRs in Reaktion auf die fremde Stimme negativer aus als die der MMRs, welche durch die mütterliche Stimme hervorgerufen wurden. Im späten Zeitfenster kann die FA als Reaktion auf die mütterliche Stimme als tendenziell positiver beschrieben werden als die der MMRs, die durch die fremde Stimme hervorgerufen wurde – jedoch ohne Signifikanz zu erreichen.

Unsere Ergebnisse deuten darauf hin, dass eine fremde Stimme eine frühe negative Komponente (FN) bei Säuglingen auslöst, die der MMN im Erwachsenenalter ähnelt. Die FN könnte einen präattentiven Prozess darstellen, auditive Abweichungen zu detektieren. Die Abwesenheit einer ähnlichen Komponente in den MMRs der mütterlichen Stimme könnte mit dem Phänomen erklärt werden, dass die im späteren Zeitfenster gefundene positive Komponente die FN überlappte und somit maskierte. Unsere Daten legen nahe, dass die späte PK ein Korrelat für Prozesse darstellt, die hinter der Verlagerung von Aufmerksamkeit in Richtung der bekannten mütterlichen Stimme liegen. Diese Verlagerung in Richtung der mütterlichen Stimme lässt vermuten, dass diese einen Stimulus von hoher Relevanz in der frühen Kindheit repräsentiert. Wie von einigen Autoren bereits als Hypothese aufgestellt, könnte die PK somit ein Vorläufer der erwachsenen P3a Komponente sein. Zudem interpretierten wir das Hervorrufen einer PK durch die mütterliche Stimme als Ausdruck, dass Prozesse, die dem auditiven Wiedererkennungsgedächtnis unterliegen präattentiven Prozesse beeinflussen, Abweichungen in einer Folge von regelmäßigen Stimuli zu detektieren.

Aus unseren Ergebnissen lässt sich der Schluss ziehen, dass Säuglinge, die nach der Geburt über einen längeren Zeitraum Stimmerfahrungen sammeln konnten, die mütterliche Stimme neuronal anders verarbeiten, als eine unbekannte, weibliche Stimme. Unsere Ergebnisse bekräftigen zudem die Beobachtungen von Beauchemin und Kollegen: Ein hochgradig bekannter Stimmstimulus löst bei Neugeborenen eine positive MMR aus (Beauchemin et al., 2011).

Darüber hinaus zeigen unsere Beobachtungen, dass sich eine robuste neuronale Antwort auf auditive Abweichungen wahrscheinlich in den ersten Lebensmonaten bzw. zu einem Zeitpunkt, an dem der Säugling ein gewisses Maß an Stimmerfahrung gesammelt hat etabliert. Unsere Ergebnisse bekräftigen die Theorie, dass ein Teil der neuronalen Stimmverarbeitung der frühen Kindheit präattentiv abläuft und die ihr unterliegenden Prozesse zum Teil durch Stimmvorlagen des Langzeitgedächtnisses moduliert werden (Beauchemin et al., 2006; Beauchemin et al., 2011).

8 LIST OF REFERENCES

Abrams, D. A., Chen, T., Odriozola, P., Cheng, K. M., Baker, A. E., Padmanabhan, A., . . . Menon, V. (2016). Neural circuits underlying mother's voice perception predict social communication abilities in children. *Proceedings of the National Academy of Sciences of the United States of America, 113*(22), 6295-6300. doi:10.1073/pnas.1602948113

Albrecht, R., Suchodoletz, W., & Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clinical Neurophysiology, 111*(12), 2268-2276.

Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., & Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology, 77*(2), 151-155.

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jaaskelainen, I. P., . . . Ilmoniemi, R. J. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology, 35*(2), 211-224.

Apgar, V. (1953). A proposal for a new method of evaluation of the newborn infant. *Current Researches in Anesthesia & Analgesia, 32*(4), 260-267.

Atienza, M., & Cantero, J. L. (2001). Complex sound processing during human REM sleep by recovering information from long-term memory as revealed by the mismatch negativity (MMN). *Brain Research, 901*(1-2), 151-160.

Baldeweg, T., Williams, J. D., & Gruzelier, J. H. (1999). Differential changes in frontal and sub-temporal components of mismatch negativity. *International Journal of Psychophysiology, 33*(2), 143-148.

Beauchemin, M., De Beaumont, L., Vannasing, P., Turcotte, A., Arcand, C., Belin, P., & Lassonde, M. (2006). Electrophysiological markers of voice familiarity. *European Journal of Neuroscience, 23*(11), 3081-3086. doi:10.1111/j.1460-9568.2006.04856.x

Beauchemin, M., Gonzalez-Frankenberger, B., Tremblay, J., Vannasing, P., Martinez-Montes, E., Belin, P., . . . Lassonde, M. (2011). Mother and stranger: an electrophysiological study of voice processing in newborns. *Cerebral Cortex, 21*(8), 1705-1711. doi:10.1093/cercor/bhq242

Belin, P., Bestelmeyer, P. E., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology, 102*(4), 711-725. doi:10.1111/j.2044-8295.2011.02041.x

Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences, 8*(3), 129-135. doi:10.1016/j.tics.2004.01.008

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(6767), 309-312. doi:10.1038/35002078

Brannon, E. M., Libertus, M. E., Meck, W. H., & Woldorff, M. G. (2008). Electrophysiological measures of time processing in infant and adult brains: Weber's Law holds. *Journal of Cognitive Neuroscience, 20*(2), 193-203. doi:10.1162/jocn.2008.20016

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77 (Pt 3)*, 305-327.

Capilla, A., Belin, P., & Gross, J. (2013). The early spatio-temporal correlates and task independence of cerebral voice processing studied with MEG. *Cerebral Cortex, 23*(6), 1388-1395. doi:10.1093/cercor/bhs119

Ceponiene, R., Lepisto, T., Soininen, M., Aronen, E., Alku, P., & Näätänen, R. (2004). Event-related potentials associated with sound discrimination versus novelty detection in children. *Psychophysiology, 41*(1), 130-141. doi:10.1111/j.1469-8986.2003.00138.x

Charest, I., Pernet, C. R., Rousselet, G. A., Quinones, I., Latinus, M., Fillion-Bilodeau, S., . . . Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience, 10*, 127. doi:10.1186/1471-2202-10-127

Charrier, I., Mathevon, N., & Jouventin, P. (2001). Mother's voice recognition by seal pups. *Nature, 412*(6850), 873. doi:10.1038/35091136

Cheng, Y., Lee, S.-Y., Chen, H.-Y., Wang, P.-Y., & Decety, J. (2012). Voice and Emotion Processing in the Human Neonatal Brain. *Journal of Cognitive Neuroscience, 24*(6), 1411-1419. doi:10.1162/jocn_a_00214

Cheour, M. (2007). Development of mismatch negativity (MMN) during infancy. In H. Michelle de (Ed.), *Infant EEG and event-related potentials* (pp. 171-199): Psychology Press.

Cheour, M., Alho, K., Ceponiene, R., Reinikainen, K., Sainio, K., Pohjavuori, M., . . . Näätänen, R. (1998). Maturation of mismatch negativity in infants. *International Journal of Psychophysiology, 29*(2), 217-226.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*(5), 351-353.

Cheour, M., Ceponiene, R., Leppanen, P., Alho, K., Kujala, T., Renlund, M., . . . Näätänen, R. (2002). The auditory sensory memory trace decays rapidly in newborns. *Scandinavian Journal of Psychology, 43*(1), 33-39.

Cheour, M., Leppanen, P. H., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology, 111*(1), 4-16.

Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpaa, M., Kero, P., . . . Hamalainen, H. (2002). Speech sounds learned by sleeping newborns. *Nature, 415*(6872), 599-600. doi:10.1038/415599b

Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., . . . Näätänen, R. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research, 82*(1), 53-58.

Cheour-Luhtanen, M., Alho, K., Sainio, K., Rinne, T., Reinikainen, K., Pohjavuori, M., . . . Näätänen, R. (1996). The ontogenetically earliest discriminative response of the human brain. *Psychophysiology, 33*(4), 478-481.

Coch, D., & Gullick, M. M. (2012). Event-Related Potentials and Development. In S. J. Luck (Ed.), *The Oxford handbook of event-related potential components* (pp. 475-512). Oxford: Oxford Univ. Press.

Courchesne, E., Hillyard, S. A., & Galambos, R. (1975). Stimulus novelty, task relevance and the visual evoked potential in man. *Electroencephalography and Clinical Neurophysiology, 39*(2), 131-143.

de Haan, M. (2007). Current and future directions in infant electrophysiology In M. de Haan (Ed.), *Infant EEG and event-related potentials* (pp. 305-3016). New York: Psychology Press.

DeBoer, T., Scott, L. S., & Nelson, C. A. (2007). Methods for acquiring and analyzing infant event-related potentials. In H. Michelle de (Ed.), *Infant EEG and event-related potentials* (pp. 5-38): Psychology Press.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176.

DeCasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: preference, discrimination, and reinforcing value. *Developmental Psychobiology, 17*(5), 481-491. doi:10.1002/dev.420170506

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport, 8*(4), 919-924.

Dehaene-Lambertz, G., Montavont, A., Jobert, A., Allirol, L., Dubois, J., Hertz-Pannier, L., & Dehaene, S. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language, 114*(2), 53-65. doi:10.1016/j.bandl.2009.09.003

Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport, 12*(14), 3155-3158.

deRegnier, R. A. (2007). Auditory recognition memory in infancy. In H. Michelle de (Ed.), *Infant EEG and event-related potentials* (1 ed., pp. 145-165): Psychology Press.

deRegnier, R. A., Long, J. D., Georgieff, M. K., & Nelson, C. A. (2007). Using event-related potentials to study perinatal nutrition and brain development in infants of diabetic mothers. *Developmental Neuropsychology, 31*(3), 379-396. doi:10.1080/87565640701229524

deRegnier, R. A., Nelson, C. A., Thomas, K. M., Wewerka, S., & Georgieff, M. K. (2000). Neurophysiologic evaluation of auditory recognition memory in healthy newborn infants and infants of diabetic mothers. *Journal of Pediatrics, 137*(6), 777-784. doi:10.1067/mpd.2000.109149

deRegnier, R. A., Wewerka, S., Georgieff, M. K., Mattia, F., & Nelson, C. A. (2002). Influences of postconceptional age and postnatal experience on the development of auditory recognition memory in the newborn infant. *Developmental Psychobiology, 41*(3), 216-225. doi:10.1002/dev.10070

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*(3), 357-427. doi:10.1017/S0140525X00058027

Eggermont, J. J. (1992). Development of auditory evoked potentials. *Acta Oto-Laryngologica, 112*(2), 197-200.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience, 10*(5), 590-604.

Fenton, T. R., & Kim, J. H. (2013). A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatrics, 13*, 59. doi:10.1186/1471-2431-13-59

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*(2), 181-195. doi:https://doi.org/10.1016/S0163-6383(85)80005-9

Fifer, W. P., & Moon, C. M. (1994). The role of mother's voice in the organization of brain function in the newborn. *Acta Pædiatrica, 83*, 86-93. doi:10.1111/j.1651-2227.1994.tb13270.x

Fifer, W. P., & Moon, C. M. (1995). The effects of fetal experience with sound. In J. P. Lecanuet, W. P. Fifer, N. A. Krasnegor, & W. P. Smotherman (Eds.), *Fetal development: A psychobiological perspective* (pp. 351-366). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Frank, D. W., Yee, R. B., & Polich, J. (2012). P3a from white noise. *International Journal of Psychophysiology, 85*(2), 236-241. doi:10.1016/j.ijpsycho.2012.04.005

Friederici, A. D., Friedrich, M., & Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *Neuroreport, 13*(10), 1251-1254.

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews, 25*(4), 355-373. doi:https://doi.org/10.1016/S0149-7634(01)00019-7

Fujioka, T., Mourad, N., He, C., & Trainor, L. J. (2011). Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. *Clinical Neurophysiology, 122*(1), 43-51. doi:10.1016/j.clinph.2010.04.036

Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron, 65*(6), 852-858. doi:10.1016/j.neuron.2010.03.001

Háden, G. P., Németh, R., Török, M., & Winkler, I. (2016). Mismatch response (MMR) in neonates: Beyond refractoriness. *Biological Psychology, 117*, 26-31. doi:http://dx.doi.org/10.1016/j.biopsycho.2016.02.004

He, C., Hotson, L., & Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience, 19*(5), 878-892. doi:10.1162/jocn.2007.19.5.878

He, C., Hotson, L., & Trainor, L. J. (2009). Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. *European Journal of Neuroscience, 29*(4), 861-867. doi:10.1111/j.1460-9568.2009.06625.x

Hillyard, S. A., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34*, 33-61. doi:10.1146/annurev.ps.34.020183.000341

Hirasawa, K., Kurihara, M., & Konishi, Y. (2002). The relationship between mismatch negativity and arousal level. Can mismatch negativity be an index for evaluating the arousal level in infants? *Sleep Medicine, 3 Suppl 2*, S45-48.

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology, 37*(3), 187-209. doi:10.1080/87565641.2011.627958

Hoffmann, S., & Falkenstein, M. (2008). The correction of eye blink artefacts in the EEG: a comparison of two prominent methods. *PloS One, 3*(8), e3004. doi:10.1371/journal.pone.0003004

Huotilainen, M., Kujala, A., & Alku, P. (2001). Long-term memory traces facilitate short-term memory trace formation in audition in humans. *Neuroscience Letters, 310*(2-3), 133-136.

Huotilainen, M., Kujala, A., Hotakainen, M., Parkkonen, L., Taulu, S., Simola, J., . . . Näätänen, R. (2005). Short-term memory functions of the human fetus recorded with magnetoencephalography. *Neuroreport, 16*(1), 81-84.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*(2), 167-178.

Huttenlocher, P. R., De Courten, C., Garey, L. J., & Van der Loos, H. (1982). Synaptic development in human cerebral cortex. *International Journal of Neurology, 16-17*, 144-154.

Jabès, A., & Nelson, C. A. (2015). 20 years after “The ontogeny of human memory: A cognitive neuroscience perspective,” where are we? *International Journal of Behavioral Development, 39*(4), 293-303. doi:10.1177/0165025415575766

Jing, H., & Benasich, A. A. (2006). Brain responses to tonal changes in the first two years of life. *Brain and Development, 28*(4), 247-256. doi:10.1016/j.braindev.2005.09.002

Johnson, M. H., de Haan, M., Oliver, A., Smith, W., Hatzakis, H., Tucker, L. A., & Csibra, G. (2001). Recording and analyzing high-density event-related potentials with infants. Using the Geodesic sensor net. *Developmental Neuropsychology, 19*(3), 295-323. doi:10.1207/S15326942DN1903_4

Kekoni, J., Hamalainen, H., Saarinen, M., Grohn, J., Reinikainen, K., Lehtokoski, A., & Naatanen, R. (1997). Rate effect and mismatch responses in the somatosensory system: ERP-recordings in humans. *Biological Psychology, 46*(2), 125-142.

Kisilevsky, B. S., Hains, S. M., Lee, K., Xie, X., Huang, H., Ye, H. H., . . . Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science, 14*(3), 220-224. doi:10.1111/1467-9280.02435

Kisilvesky, B. S., & Muir, D. W. (1991). Human fetal and subsequent newborn responses to sound and vibration. *Infant Behavior and Development, 14*(1), 1-26. doi:http://dx.doi.org/10.1016/0163-6383(91)90051-S

Klem, G. H., Luders, H. O., Jasper, H. H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology. Supplement, 52*, 3-6.

Kujala, T., Tervaniemi, M., & Schroger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biological Psychology, 74*(1), 1-19. doi:10.1016/j.biopsycho.2006.06.001

Kurtzberg, D., Hilpert, P. L., Kreuzer, J. A., & Vaughan, H. G., Jr. (1984). Differential maturation of cortical auditory evoked potentials to speech sounds in normal fullterm and very low-birthweight infants. *Developmental Medicine and Child Neurology, 26*(4), 466-475.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilaine, M., & Naatane, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Neuroreport, 13*(1), 47-51.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., & Näätänen, R. (2002). Maturation of the auditory change detection response in infants: a longitudinal ERP study. *Neuroreport, 13*(15), 1843-1848.

Kushnerenko, E., Van den Bergh, B. R., & Winkler, I. (2013). Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. *Frontiers in Psychology, 4*, 595. doi:10.3389/fpsyg.2013.00595

Kushnerenko, E., Winkler, I., Horvath, J., Näätänen, R., Pavlov, I., Fellman, V., & Huotilainen, M. (2007). Processing acoustic change and novelty in newborn infants. *European Journal of Neuroscience, 26*(1), 265-274. doi:10.1111/j.1460-9568.2007.05628.x

Lecanuet, J. P., & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics, Gynecology, and Reproductive Biology, 68*(1-2), 1-23.

Leppanen, P. H., Guttorm, T. K., Pihko, E., Takkinen, S., Eklund, K. M., & Lyytinen, H. (2004). Maturational effects on newborn ERPs measured in the mismatch negativity paradigm. *Experimental Neurology, 190 Suppl 1*, S91-101. doi:10.1016/j.expneurol.2004.06.002

Levy, D. A., Granot, R., & Bentin, S. (2001). Processing specificity for human voice stimuli: electrophysiological evidence. *Neuroreport, 12*(12), 2653-2657.

Luaute, J., Fischer, C., Adeleine, P., Morlet, D., Tell, L., & Boisson, D. (2005). Late auditory and event-related potentials can be useful to predict good functional outcome after coma. *Archives of Physical Medicine and Rehabilitation, 86*(5), 917-923. doi:10.1016/j.apmr.2004.08.011

Luck, S. J. (2014a). Comparing Conditions with Different Numbers of Trials or Differnet Levels of Noise In S. J. Luck (Ed.), An Introduction to the Event-related Potential Technique (2 ed.): The MIT Press. Retrieved from https://mitpress.mit.edu/sites/default/files/8575_009a_0.pdf.

Luck, S. J. (2014b). *An Introduction to the Event-related Potential Technique* (S. J. Luck Ed. 2 ed.). Massachusetts, United States: MIT Press Ltd.

Mai, X., Xu, L., Li, M., Shao, J., Zhao, Z., deRegnier, R. A., . . . Lozoff, B. (2012). Auditory recognition memory in 2-month-old infants as assessed by event-related potentials. *Developmental Neuropsychology, 37*(5), 400-414. doi:10.1080/87565641.2011.650807

Martynova, O., Kirjavainen, J., & Cheour, M. (2003). Mismatch negativity and late discriminative negativity in sleeping human newborns. *Neuroscience Letters, 340*(2), 75-78.

Moon, C., Zernzach, R. C., & Kuhl, P. K. (2015). Mothers say “baby” and their newborns do not choose to listen: a behavioral preference study to compare with ERP results. *Frontiers in Human Neuroscience, 9*, 153. doi:10.3389/fnhum.2015.00153

Näätänen, R. (2000). Mismatch negativity (MMN): perspectives for application. *International Journal of Psychophysiology, 37*(1), 3-10. doi:http://dx.doi.org/10.1016/S0167-8760(00)00091-X

Näätänen, R. (2012). The Mismatch Negativity (MMN). In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 143-157): Oxford Universtity Press.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., . . . Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*(6615), 432-434. doi:10.1038/385432a0

Näätänen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. *Biological Psychology, 8*(2), 81-136.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425.

Näätänen, R., Teder, W., Alho, K., & Lavikainen, J. (1992). Auditory attention and selective input modulation: a topographical ERP study. *Neuroreport, 3*(6), 493-496.

Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research, 20*(1), 26-36. doi:https://doi.org/10.1016/j.cogbrainres.2003.12.011

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schroger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage, 15*(1), 167-174. doi:10.1006/nimg.2001.0970

Otte, R. A., Winkler, I., Braeken, M. A., Stekelenburg, J. J., van der Stelt, O., & Van den Bergh, B. R. (2013). Detecting violations of temporal regularities in waking and sleeping two-month-old infants. *Biological Psychology, 92*(2), 315-322. doi:10.1016/j.biopsycho.2012.09.009

Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences of the United States of America, 110*(37), 15145-15150. doi:10.1073/pnas.1302159110

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., . . . Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*(2), 127-152. doi:10.1111/1469-8986.3720127

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology, 111*(2), 220-236.

Pratt, H. (2012). Sensory ERP components. In S. J. Luck (Ed.), *The Oxford handbook of event-related potential components* (pp. 89-114). Oxford: Oxford University Press.

Purhonen, M., Kilpelainen-Lees, R., Valkonen-Korhonen, M., Karhu, J., & Lehtonen, J. (2005). Four-month-old infants process own mother's voice faster than unfamiliar voices--electrical signs of sensitization in infant brain. *Brain Research: Cognitive Brain Research, 24*(3), 627-633. doi:10.1016/j.cogbrainres.2005.03.012

Querleu, D., Lefebvre, C., Titran, M., Renard, X., Morillion, M., & Crepin, G. (1984). Reaction of the newborn infant less than 2 hours after birth to the maternal voice. *Journal de Gynécologie, Obstétrique et Biologie de la Reproduction, 13*(2), 125-134.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crepin, G. (1988). Fetal hearing. *European Journal of Obstetrics, Gynecology, and Reproductive Biology, 28*(3), 191-212.

Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Vervoort, P. (1988). [Intra-amniotic transmission of the human voice]. *Revue Française de Gynécologie et d'Obstétrique, 83*(1), 43-50.

Rogier, O., Roux, S., Belin, P., Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4- to 5-year-old children. *International Journal of Psychophysiology, 75*(1), 44-47. doi:10.1016/j.ijpsycho.2009.10.013

Sadeh, A. (2004). A brief screening questionnaire for infant sleep problems: validation and findings for an Internet sample. *Pediatrics, 113*(6), e570-577.

Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The Human Auditory Sensory Memory Trace Persists about 10 sec: Neuromagnetic Evidence. *Journal of Cognitive Neuroscience, 5*(3), 363-370. doi:10.1162/jocn.1993.5.3.363

Scott, S. K. (2008). Voice processing in monkey and human brains. *Trends in Cognitive Sciences, 12*(9), 323-325. doi:10.1016/j.tics.2008.06.003

Seress, L., Abraham, H., Tornoczky, T., & Kosztolanyi, G. (2001). Cell formation in the human hippocampal formation from mid-gestation to the late postnatal period. *Neuroscience, 105*(4), 831-843.

Siddappa, A. M., Georgieff, M. K., Wewerka, S., Worwa, C., Nelson, C. A., & Deregnier, R. A. (2004). Iron deficiency alters auditory recognition memory in newborn infants of diabetic mothers. *Pediatric Research, 55*(6), 1034-1041. doi:10.1203/01.pdr.0000127021.38207.62

Squire, L. R., Schmolck, H., & Stark, S. M. (2001). Impaired auditory recognition memory in amnesic patients with medial temporal lobe lesions. *Learning and Memory, 8*(5), 252-256. doi:10.1101/lm.42001

Stahl, D., Parise, E., Hoehl, S., & Striano, T. (2010). Eye contact and emotional face processing in 6-month-old infants: Advanced statistical methods applied to event-related potentials. *Brain and Development, 32*(4), 305-317. doi:http://dx.doi.org/10.1016/j.braindev.2009.04.001

Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch negativity in the visual modality. *Neuroreport, 10*(16), 3363-3367.

Therien, J. M., Worwa, C. T., Mattia, F. R., & deRegnier, R. A. (2004). Altered pathways for auditory discrimination and recognition memory in preterm infants. *Developmental Medicine and Child Neurology, 46*(12), 816-824.

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature, 372*(6501), 90-92. doi:10.1038/372090a0

Titova, N., & Naatanen, R. (2001). Preattentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neuroscience Letters, 308*(1), 63-65.

Trainor, L. (2007). Event-Related Potential (ERP) Measures in Auditory Development Research. In L. A. Schmidt & S. J. Segalowitz (Eds.), *Developmental Psychophysiology: Theory, Systems, and Methods* (pp. 69-102). New York: Cambridge University Press.

Trainor, L. (2008). Science & music: the neural roots of music. *Nature, 453*(7195), 598-599. doi:10.1038/453598a

Trainor, L., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology, 51*(1), 5-15.

Van Lancker, D. R., & Canter, G. J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition, 1*(2), 185-195.

Van Lancker, D. R., Cummings, J. L., Kreiman, J., & Dobkin, B. H. (1988). Phonagnosia: a dissociation between familiar and unfamiliar voices. *Cortex, 24*(2), 195-209.

Van Lancker, D. R., & Kreiman, J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologia, 25*(5), 829-834.

van Leeuwen, T., Been, P., Kuijpers, C., Zwarts, F., Maassen, B., & van der Leij, A. (2006). Mismatch response is absent in 2-month-old infants at risk for dyslexia. *Neuroreport, 17*(4), 351-355. doi:10.1097/01.wnr.0000203624.02082.2d

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research, 742*(1-2), 239-252.

Wunderlich, J. L., Cone-Wesson, B. K., & Shepherd, R. (2006). Maturation of the cortical auditory evoked potential in infants and young children. *Hearing Research, 212*(1–2), 185-202. doi:http://doi.org/10.1016/j.heares.2005.11.010

Zinke, K., Thone, L., Bolinger, E. M., & Born, J. (2018). Dissociating Long and Short-term Memory in Three-Month-Old Infants Using the Mismatch Response to Voice Stimuli. *Frontiers in Psychology, 9*, 31. doi:10.3389/fpsyg.2018.00031

9 ERKLÄRUNG ZUM EIGENANTEIL

Die Arbeit wurde am Institut für Medizinische Psychologie und Verhaltensneurobiologie in Tübingen unter Betreuung von Prof. Dr. Jan Born durchgeführt.

Die Konzeption der Studie erfolgte durch Dr. Katharina Zinke, Mitarbeiterin des Institutes für Medizinische Psychologie und Verhaltensneurobiologie.

Die Versuche wurden nach Einarbeitung von Dr. Katharina Zinke von mir mit Unterstützung von Gaby Beck und Astrid Günther durchgeführt. Die Vorverarbeitung der EEG-Daten (Filtern, Schneiden, Artefaktentfernung) wurde größtenteils von Dr. Katharina Zinke übernommen. Die Weiterverarbeitung und Quantifizierung wurden nach Anleitung durch Dr. Katharina Zinke durchgeführt.

Die statistische Auswertung erfolgte nach Anleitung durch Dr. Katharina Zinke durch mich.

Ich versichere, das Manuskript inklusive Abbildungen eigenständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

Tübingen, den

10 PUBLICATIONS

Zinke, K., Thone, L., Bolinger, E. M., & Born, J. (2018). Dissociating Long and Short-term Memory in Three-Month-Old Infants Using the Mismatch Response to Voice Stimuli. *Frontiers in Psychology, 9*, 31. doi:10.3389/fpsyg.2018.0003

11 ACKNOWLEDGEMENT

It would not have been possible to write my doctoral thesis without the support of the team of the Institute of Medical Psychology and Behavioral Neurobiology in Tübingen. First, I would like to thank my supervisor Katharina Zinke for her ideas on the study setting, her support when analyzing our data and for giving structure but and at the same time autonomy when discussing problems that we encountered in the process of the study. I am grateful that I had the opportunity to work in such an inspiring and motivating environment. Furthermore, I sincerely thank for getting an insight into neurocognitive basic research, a field that had been uncharted waters for me at the beginning, since I studied medicine.

Furthermore, I would like to thank Jan Born for his direction-giving suggestions on the study setting and data analysis.

I warmly thank Astrid Günther and Gaby Beck for their technical and emotional support when receiving and handling the little participants and their parents during the test sessions. Finally, I express my gratitude to the 35 families that participated with their sons and daughters in the study, for their open-mindedness as well as patience during the testing sessions.

12 CURRICULUM VITAE

Personal Information:

Name: Leonie Thöne

Date of Birth: 11.07.1990

Place of Birth: Göttingen

Nationality: German

Contact: leonie.thoene@gmail.com

Education:

2010-2016 Medicine at the University of Tübingen and at the University of Coimbra, Portugal

(2. Staatsexamen: 15.10.2015, 3. Staatsexamen: 22.11.2016)

2013-2017 Doctoral student at the Institute of Medical Psychology and Behavioral Neurobiology, Tübingen

Professional Experience:

2015/2016 Practical year as medical student:

1. Department of Internal Medicine, Katharinenhospital, Stuttgart
2. Department of Neonatology and Paediatrics, Inselspital Bern, Switzerland
3. Department of Paediatric and General Surgery, Hospital Universitário, Cartagena, Colombia

Current position: Assistant doctor, Department of Paediatrics, University Medicine of Göttingen

13 APPENDIX

VP Nr.	Untersucher	Datum
Angaben zum Kind		
Geschlecht	Alter in Tagen	Geburtsdatum
Schwangerschaftsalter (Schwangerschaftswoche) bei Geburt		
Gab es irgendwelche Komplikationen bei der Geburt?		
Gewicht und Größe bei Geburt	Kopfumfang bei Geburt	Apgar Score
Gewicht und Größe aktuell	Kopfumfang aktuell	
Hat Ihr Kind Auffälligkeiten? Wenn ja, welche?		
Wurde bei Ihrem Kind ein Hörtest gemacht? wenn ja, was war das Ergebnis?		
Die folgenden Fragen beziehen sich auf den letzten Monat		
Wo schläft Ihr Kind in der Regel? ☐ Kinderbett im separatem Raum ☐ Kinderbett im Raum der Eltern ☐ Im Bett der Eltern ☐ Kinderbett im Raum mit Geschwistern ☐ Sonstiges :		
Schläft Ihr Kind vorwiegend in ☐ Bauchlage ☐ Seitenlage oder ☐ Rückenlage?		
Wie viel Zeit verbringt Ihr Kind schlafend am Tag (zwischen 7 Uhr morgens und 7 Uhr abends)?		
Durchschnittliche Zahl des Aufwachens pro Nacht?		
Wie viel Zeit in der Nacht verbringt Ihr Kind wach (von 10 Uhr abends bis 6 Uhr morgens)?		
Wie lange dauert es, Ihr Kind zum Einschlafen zu bringen?		

Wie schläft Ihr Kind ein? ☐ Während des Fütterns, ☐ wenn es gewiegt wird, ☐ Wenn es gehalten wird, ☐ Allein im Bett, ☐ Im Bett, nahe eines Elternteils	
Wann schläft Ihr Baby normalerweise abends ein? Stunde (z.B. 19): _______ Minute (z.B. 20)________	
Betrachten Sie den Schlaf Ihres Kindes als Problem? ☐ Ein sehr ernstes Problem ☐ Ein kleines Problem ☐ Überhaupt kein Problem	
Wann und wie lange hat Ihr Kind in den letzten 24 Stunden geschlafen (ungefähr)?	
Gab es in den letzten 24 Stunden Abweichungen von den Gewohnheiten Ihres Kindes?	
Welche Sprachen werden bei Ihnen zu Hause gesprochen? Welche mit dem Kind?	
Angaben zur Mutter	
Geburtsdatum	Höchster Schul-/Bildungsabschluss
Haben/Hatten Sie Probleme in dieser Schwangerschaft, welche und wann?	
Wie oft waren sie bisher schwanger - diese SS mit einbezogen? Geburten: Anzahl: ___________ Wann:___________	
Stillen Sie?	
Wann haben Sie das letzte Mal gestillt?	
Nehmen Sie Medikamente, wenn ja welche:	
Rauchern Sie, wenn ja wieviel pro Tag?	
Trinken Sie Alkohol, wie viele Gläser/ was?	