

Why do humans have unique auditory event-related fields? Evidence from computational modeling and MEG experiments

Aida Hajizadeh¹ | Artur Matysiak¹  | André Brechmann² | Reinhard König¹ | Patrick J. C. May^{1,3}

¹Leibniz Institute for Neurobiology, Research Group Comparative Neuroscience, Magdeburg, Germany

²Leibniz Institute for Neurobiology, Combinatorial NeuroImaging Core Facility, Magdeburg, Germany

³Department of Psychology, Lancaster University, Lancaster, UK

Correspondence

Artur Matysiak, Leibniz Institute for Neurobiology, Research Group Comparative Neuroscience, Brenneckestraße 6, 39118 Magdeburg, Germany.
Email: Artur.Matysiak@lin-magdeburg.de

Funding information

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 763959

Abstract

Auditory event-related fields (ERFs) measured with magnetoencephalography (MEG) are useful for studying the neuronal underpinnings of auditory cognition in human cortex. They have a highly subject-specific morphology, albeit certain characteristic deflections (e.g., P1m, N1m, and P2m) can be identified in most subjects. Here, we explore the reason for this subject-specificity through a combination of MEG measurements and computational modeling of auditory cortex. We test whether ERF subject-specificity can predominantly be explained in terms of each subject having an individual cortical gross anatomy, which modulates the MEG signal, or whether individual cortical dynamics is also at play. To our knowledge, this is the first time that tools to address this question are being presented. The effects of anatomical and dynamical variation on the MEG signal is simulated in a model describing the core-belt-parabelt structure of the auditory cortex, and with the dynamics based on the leaky-integrator neuron model. The experimental and simulated ERFs are characterized in terms of the N1m amplitude, latency, and width. Also, we examine the waveform grand-averaged across subjects, and the standard deviation of this grand average. The results show that the intersubject variability of the ERF arises out of both the anatomy and the dynamics of auditory cortex being specific to each subject. Moreover, our results suggest that the latency variation of the N1m is largely related to subject-specific dynamics. The findings are discussed in terms of how learning, plasticity, and sound detection are reflected in the auditory ERFs. The notion of the grand-averaged ERF is critically evaluated.

KEYWORDS

anatomy, auditory cortex, computational modeling, dynamics, event-related field, ERF, latency, magnetoencephalography, MEG, N1m

1 | INTRODUCTION

The auditory event-related response is revealed by presenting a stimulus multiple times, and then, averaging the evoked

magnetoencephalography (MEG) or electroencephalography (EEG) signal across the stimulus presentations. The resulting typical trial-averaged response is characterized by a sequence of peaks and troughs. The initial cortically generated ones,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research

called long-latency responses, reach extrema at approximately 50, 100, and 200 ms after stimulus onset. The respective labels for these responses are P1, N1, and P2 when these are observed in EEG as the part of the event-related potential (ERP). The corresponding labels are P1m, N1m, and P2m when measurements are done in MEG to reveal the event-related field (ERF). Of these, the N1/N1m tends to be the most prominent response. However, there is a large between-subject variability of auditory event-related responses. The peak amplitude of the N1m ranges from several tens of fT to almost 1 pT, a similar order-of-magnitude variation can be seen with the N1, and the peak latency of the N1/N1m is observed in the 70–130 ms range. Thus, averaging the peak amplitude and the peak latency across subjects results in sizeable standard deviations. Also, the grand-averaged response tends to be broader than any of the individual responses because of the variations in peak latencies. Further, the waveform of the event-related response comes in many morphological varieties. The N1/N1m can have a double-peak structure, or its rising or falling slope can have a pronounced shoulder, and many subjects exhibit no clear P2/P2m response at all.

Importantly, the intersubject variability of the event-related response is not merely due to noise in the measurement. Rather, ERFs and ERPs are intrinsically subject-specific, remaining stable from measurement session to session, when these are separated by days, weeks, or even years (see, e.g., Ahonen et al., 2016; Atcherson et al., 2006; Dalebout & Robey, 1997; Michalewski et al., 1986; Sandman & Patterson, 2000; Segalowitz & Barnes, 1993). Figure 1 demonstrates the intrasubject stability (reproducibility) and the intersubject variability of auditory ERFs recorded from two subjects in our laboratory. The test–retest measurements were performed with the identical experimental paradigm and auditory stimuli (simple tone repetition), and the interval between the two MEG recordings was a year for one subject and 3 years for the other. For each subject, the two recordings are from the same MEG channel above the temporal lobe that shows the largest N1m response. There are two observations to be made. First, the two subjects produce substantially different waveforms, with subject-specific peak amplitudes and latencies. Second, the waveforms are reproducible across the long time intervals.

One interpretation of the between-subject differences in auditory ERFs relates to anatomical variations of the auditory cortex (AC) between individual subjects and, within a subject, between the two hemispheres. This is because the magnetic field generated by source activity in cortex depends on the source's orientation and on its distance to the measuring sensor, and these in turn are determined by the topography of the cortex, that is, the ridges and folds of the cortical surface (Hämäläinen et al., 1993). For example, Shaw et al. (2013) concluded that the rightward bias of the N1m amplitude, a phenomenon frequently observed in MEG measurements, is based on a larger degree of cortical folding in the left compared to the right hemisphere. Anatomical

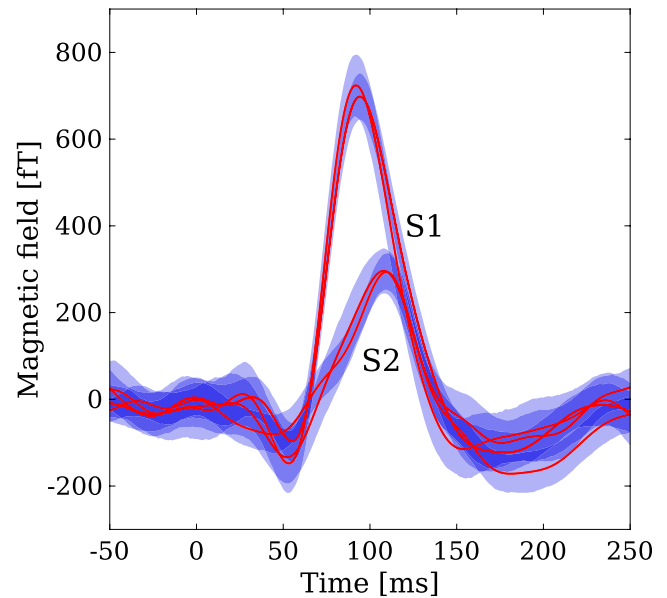


FIGURE 1 Examples of intrasubject reproducibility and subject-specificity of ERFs. The figure shows trial-averaged ERFs evoked in two subjects, S1 and S2, by a sequence of identical stimuli (1.5-kHz tone, sound-pressure level: 80 dB, stimulus onset interval: 7 s, approximately 100 stimulus repetitions). The interval between the two measurements was 3 years for S1 and 1 year for S2. Data show recordings from the MEG channels (magnetometers of the two MEG systems) with the maximum absolute N1m-peak amplitude above the posterior part of the right (S1) and left (S2) hemisphere. For each subject, we found a compelling agreement between the first and second measurements. The variability of the ERFs among subjects is clearly reflected in the large differences in the N1m-peak amplitudes and peak latencies. The waveforms of the second measurement were scaled by a factor of about 1.5 to achieve good agreement between the earlier and later record. This is because we show magnetic field responses from the sensors with the largest N1m amplitude. A slight difference in the positioning of the subject's head between the measurements will scale the response. The shaded region around each of the trial-averaged waveforms represent the 95% confidence interval (CI) of the estimated mean, which was achieved by applying 1,000 bootstrap repetitions (with replacement) to the single trials underlying the means. In each subject, there is a strong overlap of the 95% CI for the test and retest waveforms

work has demonstrated that the morphology of Heschl's gyrus (HG), which harbors the primary auditory cortex, exhibits large cross-subject variability. Different morphotypes manifest themselves in different numbers of gyri, ranging from a single HG to a common stem and a complete posterior duplication to multiple duplications, which also vary between the two hemispheres (Heschl, 1878; Morosan et al., 2001; Rademacher et al., 2001; von Economo & Horn, 1930). Moreover, larger morphological differences can be observed in higher cortical areas as compared to primary areas (Fischl et al., 2008).

Theoretically, the origin of cross-subject variability of ERFs is suggested by Maxwell's equations, which, in combination with the continuity equation, forms the mathematical

basis for the computation of the MEG signal. Assuming that the brain is a conducting volume H with constant conductivity σ , the quasistatic approximation of Maxwell's equations, where all time-derivative terms can be ignored as source terms, provides simple solutions to this so-called forward problem. Specifically, the magnetic field $\mathbf{B}(\mathbf{r})$ at a location \mathbf{r} outside the brain generated by electric currents at a location \mathbf{r}' inside the brain follows the Ampère–Laplace law (Hämäläinen et al., 1993; Mosher et al., 1999; Sarvas, 1987; Williamson & Kaufman, 1981):

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \cdot \int_H \mathbf{J}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} d\mathbf{r}'. \quad (1)$$

The neural activity is described by the total current density $\mathbf{J}(\mathbf{r}') = \mathbf{J}^P(\mathbf{r}') + \mathbf{J}^V(\mathbf{r}')$. Its first component is the primary current density $\mathbf{J}^P(\mathbf{r}')$ which describes the movement of electrical charge inside the dendritic tree. The second component is the volume current density $\mathbf{J}^V(\mathbf{r}')$ which is proportional to $\mathbf{J}^P(\mathbf{r}')$ and denotes the passive return current in extracellular space. Further, it is assumed that the magnetic permeability of tissue μ is equal to the permeability μ_0 of free space. The Ampère–Laplace law indicates that the generation of ERFs can be separated into two conceptually different components connected by their cross-product: (a) the total current density $\mathbf{J}(\mathbf{r}')$ is produced by neural activity and is thus determined by the brain as a dynamical system; (b) the term $\mathbf{r} - \mathbf{r}'$ denotes the position of the intracranial current source in relation to the extracranial position of the measurement sensor, and thus, reflects the anatomy of the brain, but not the brain dynamics. Notably, ERFs show distinct—and opposite—dependencies on these two components. $\mathbf{B}(\mathbf{r})$ increases linearly with increasing source strength and decreases nonlinearly with increasing distance from the source (Brody et al., 1973; Hämäläinen et al., 1993; Sarvas, 1987; Zhang, 1995). Thus, the ERF reflects neural dynamics via its linear relationship to source strengths, and it also reflects the gross anatomy of the brain, that is, the physical layout of the sources, via a nonlinear relationship to distance and orientation of the sources. Cross-subject variability of ERFs could, therefore, arise out of subject-specific dynamics, subject-specific anatomy, or a combination of both.

In our previous work, we addressed the impact of anatomical and dynamical contributions to the auditory ERF by using simulations of auditory cortex (Figure 10a,b in Hajizadeh et al., 2019). We found that when the modulating effect of the anatomy was varied while keeping the dynamics of the model fixed, the peak amplitude of the N1m became distributed across a wide range, whereas the peak latency of the N1m was little affected. A very different picture emerged when the simulated anatomy remained fixed but the dynamical parameters of the model were varied. In this case, both the peak amplitude and the peak latency of the N1m had a wide distribution, and these two measures were strongly correlated.

The aim of this work is to investigate why ERFs vary from subject to subject by testing the predictions of our computational model (Hajizadeh et al., 2019) in MEG measurements in human subjects. The current, largely unwritten understanding attributes the subject-specificity of ERFs mainly to well-established cross-subject differences in the gross anatomy of cortex. It remains an open question to which degree subject-specific ERFs also reflect the presence of brain dynamics that is specific to the subject. Here, we address this question by linking experimental observations from previous studies (König et al., 2015; Matysiak et al., 2013) and previously unpublished data to simulations from our computational model of auditory cortex.

2 | METHOD

2.1 | Computational model

Simulations were performed on a model of auditory cortex which was originally developed to examine the consequences of short-term synaptic plasticity on auditory processing (May et al., 2015; May & Tiitinen, 2010, 2013; Westö et al., 2016). Its basic dynamical unit is the cortical column, which is described as a pool of excitatory (pyramidal) neurons interacting with a pool of inhibitory interneurons, much as in Wilson and Cowan (1972). The dynamic equations for this interaction are those of the leaky-integrator neuron (LIN; e.g., Hopfield & Tank, 1986), whereby the time derivative of the state variable, which is equivalent to the membrane potential, is proportional to the sum of a leak term and the synaptic input currents. Each current depends linearly on the presynaptic spiking rate and the synaptic strength. Furthermore, the excitatory connections between the pyramidal neurons are modulated by a term describing short-term synaptic plasticity as in Loebel et al. (2007). The output of the LIN is the instantaneous spiking rate derived by passing the state variable through a nonlinear function. In the model, each pool of neurons is described by a single state variable and a single spiking rate representing the mean activity of the pool. Thus, each cortical column is described by a pair of ordinary differential equations, one for the pool of excitatory neurons, the other for the interneurons. In the current simulations, we also included two areas of subcortical processing: the inferior colliculus (IC) and the thalamus. As with cortical columns, we assumed that their dynamical units were interacting pools of inhibitory and excitatory neurons.

Structurally, the model mimics the AC of the macaque monkey with 13 cortical fields (Hackett et al., 2014; Kaas & Hackett, 2000). The input stage of the model represents tonotopically organized IC which feeds into a tonotopically organized thalamus. The thalamocortical input stream represents the lemniscal pathway and targets three tonotopically organized, interconnected fields of the core area, also known as the primary

auditory cortex. From there, activity spreads along multiple, topographically organized feedforward pathways to eight surrounding belt fields, which form part of the secondary auditory cortex. Each belt field, in turn, is interconnected predominantly with its nearest neighboring belt fields. Activation from the belt fields is also fed forward to two parabelt fields, also part of the secondary auditory cortex. The feedforward pathways are complemented by reciprocal, feedback pathways. Though being a vastly simplified description of the AC, the model (May et al., 2015; May & Tiitinen, 2010, 2013) is able to reproduce a variety of intra- and extracortically measured effects showing that cortical activation depends both on the incoming stimulus and the historical context of that stimulation (e.g., Brosch & Schreiner, 1997, 2000; Näätänen, 1992).

The dynamical equations of the model comprising N cortical columns (May et al., 2015; May & Tiitinen, 2010, 2013) are given by:

$$\tau_m \dot{\mathbf{u}}(t) = -\mathbf{u}(t) + S(t) \circ W_{ee} \cdot g[\mathbf{u}(t)] - W_{ei} \cdot g[\mathbf{v}(t)] + \mathbf{I}_{\text{aff},e}(t), \quad (2)$$

$$\tau_m \dot{\mathbf{v}}(t) = -\mathbf{v}(t) + W_{ie} \cdot g[\mathbf{u}(t)] - W_{ii} \cdot g[\mathbf{v}(t)] + \mathbf{I}_{\text{aff},i}(t). \quad (3)$$

Here, τ_m is the membrane time constant, and $\mathbf{u}(t) = [u_1(t), \dots, u_N(t)]$ and $\mathbf{v}(t) = [v_1(t), \dots, v_N(t)]$ are time-dependent vectors of the state variables of excitatory (index “e”) and inhibitory (index “i”) cell populations, respectively. In the current simulations, there were 208 cortical columns in total distributed over 13 cortical fields, with 16 columns per field. The IC and thalamus, each comprising 16 column-like units, were also described by the above equations. Thus, there were a total of 240 dynamical units in the model. The connections between the cell populations are mathematically expressed by the four weight matrices W_{ee} , W_{ei} , W_{ie} , and W_{ii} . The elements of W_{ee} represent excitatory-to-excitatory connections, and the elements of W_{ie} describe lateral inhibition. The matrices W_{ei} and W_{ii} have diagonal elements only and describe local, within-column connections of the inhibitory-to-excitatory and inhibitory-to-inhibitory type, respectively. Note that a connection weight describes the intensity with which two populations can interact, and it encapsulates both the average synaptic strengths as well as the density of the connections. The nonlinear function $S(t)$ represents the short-term synaptic plasticity and modifies the weights between the excitatory cell populations in the weight matrix W_{ee} at each time point with an entry-wise multiplication (expressed by the symbol “ \circ ” for the Hadamard product). The spiking rate functions $g[\mathbf{u}(t)]$ and $g[\mathbf{v}(t)]$ are sigmoid functions of the state variables, and the vectors $\mathbf{I}_{\text{aff},e}(t)$ and $\mathbf{I}_{\text{aff},i}(t)$ represent the afferent input to the excitatory and inhibitory cell populations.

Due to the nonlinearities of the functions $S(t)$, $g[\mathbf{u}(t)]$, and $g[\mathbf{v}(t)]$, Equations (2) and (3) need to be solved numerically. Thus, simulations are required to investigate how the

anatomical connectivity pattern and other model parameters shape the ERF. To gain deeper insight into the confluence of stimulation, system parameters, and cortical dynamics generating the event-related response, we recently developed a linear approximation of the model (for a full treatment, see Hajizadeh et al., 2019). This approach provides explicit solutions to the system dynamics and enables the characterization of AC activity in terms of normal modes. These are damped harmonic oscillators emerging out of the excitatory and inhibitory coupling of the cortical columns; they are described by:

$$u_d(t) = \exp(-\gamma_d t) [a_{ud} \sin(\delta_d t) + b_{ud} \cos(\delta_d t)] + c_{ud}, \quad (4)$$

$$v_d(t) = \exp(-\gamma_d t) [a_{vd} \sin(\delta_d t) + b_{vd} \cos(\delta_d t)] + c_{vd}. \quad (5)$$

Here, the decay constant γ_d and the damping frequency δ_d depend solely on the connection matrices, and the coefficients a_{ud} , a_{vd} , b_{ud} , b_{vd} , c_{ud} , and c_{vd} are functions of the connection matrices and the afferent inputs. Each normal mode depends explicitly on all parameters of the system, including the pattern of the connections between all columns. Therefore, a normal mode on its own does not represent the activity of any individual column. Instead, it should be thought of as a dynamic building block that is spread across the whole system, contributing to the activity of each column with a specific weight. Conversely, the activity of any one column represents the weighted sum of all the normal modes of the system, and, thus, is directly dependent on the anatomical structure of the AC.

In the original model (May et al., 2015; May & Tiitinen, 2013), lateral inhibition was realized by the excitatory populations making lateral connections to the inhibitory populations of neighboring columns so that W_{ie} had off-diagonal elements. In order to generate the analytical solutions, it was necessary to remove these off-diagonal elements. Lateral inhibition was included in the analytical model by introducing negative connections into W_{ee} , effectively combining the original matrices W_{ee} and W_{ie} into a matrix W_{AC} . This contained all lateral and long-range (i.e., nondiagonal) connections, both excitatory and inhibitory. In practice, W_{AC} was constructed by using Gaussians with stochastic terms to determine the connection strength as a function of the distance between the connecting columns on the tonotopic map (for details, see Hajizadeh et al., 2019).

2.2 | MEG simulation

The MEG signal was calculated by approximating the primary current in each column as being a linear function of the

synaptic inputs targeting the excitatory cell population of the column. The contribution of each input was multiplied by a connection-specific anatomical factor. These factors account for the magnetic field depending not just on the strength of the primary current, but also on the distance of the current to the MEG sensor, and on the orientation of the current (Hämäläinen et al., 1993), as seen in Equation 1. The orientation of the current not only depends on the subject-specific folding of the cortical surface, which embeds the current, but it also depends on the apical-dendrite location of the synapse driving the current and on whether the synapse is excitatory or inhibitory. Thus, the MEG signal produced by the model is the product of two mutually independent factors: (a) the dynamics of the auditory cortex, as reflected in the synaptic

inputs that the cortical columns receive, and (b) the subject-specific anatomical parameters.

The anatomical parameters, as described above, are denoted by A , which is a collection of multipliers, one per connection made onto the excitatory populations. As in Hajizadeh et al. (2019), we construct these multipliers by first defining three matrices K_1 , K_2 , and K_3 comprising multipliers according to connection type. K_1 modulates the contribution made by the excitatory connections in W_{AC} , and therefore, has the same structure as W_{ee} , which encapsulates how the 13 cortical fields are connected with each other (Figure 2a). The elements of K_1 are further divided into feedforward, feedback, and within-field (diagonal) types. The matrices K_2 and K_3 have an identical structure and modulate the contribution made by the

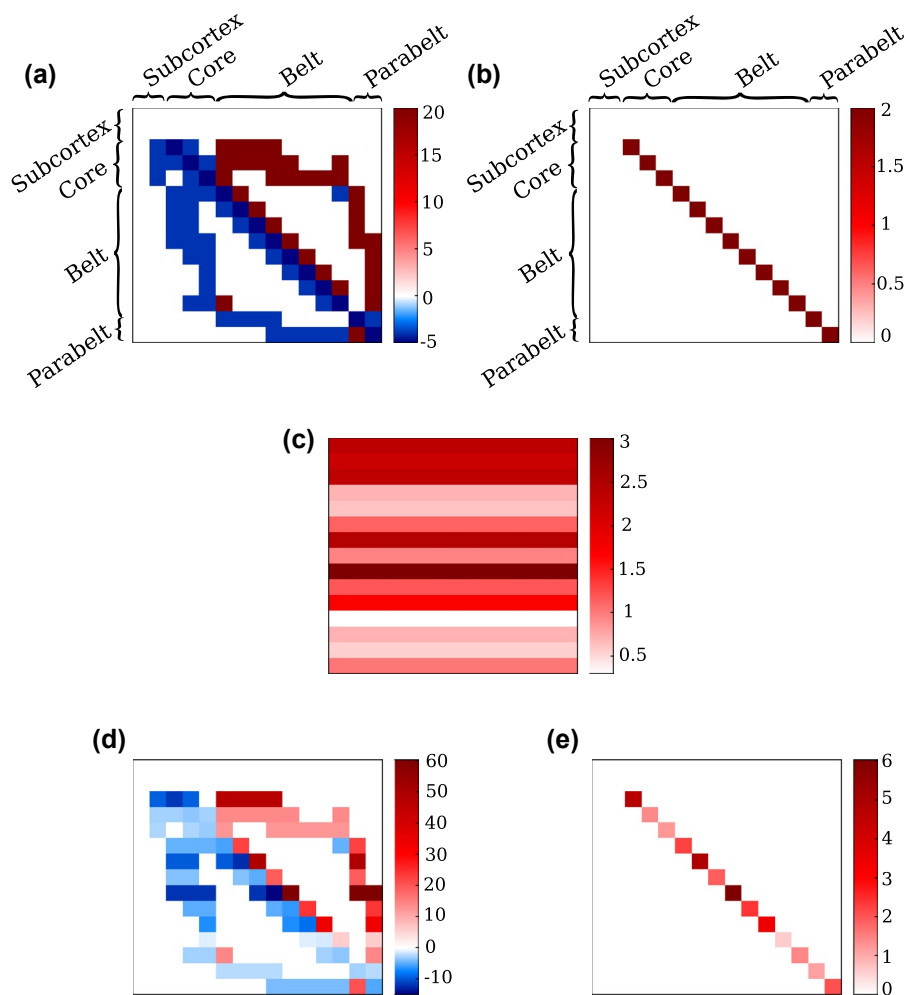


FIGURE 2 The matrices for computing the MEG signal. (a) The matrix K_1 contains the multipliers for the contribution of the MEG signal coming from the excitatory connections of type feedforward (blue), feedback (dark red), and intra-field (dark blue). (b) The matrices K_2 and K_3 are identical to each other, and they provide multipliers for the intra-column inhibitory connections and the lateral inhibitory connections, respectively. (c) The topography matrix T represents the gross anatomy of auditory cortex and it modulates the MEG in a field-specific way. Each row represents the field-specific effect that the field has on the MEG signal via orientation and distance to the sensor. (d, e) Element-wise multiplication of K_i and the T results in the final multipliers which modulate the contribution to the MEG from each connection. Note that this figure displays 15×15 matrices where the indexing runs over the 2 subcortical and 13 cortical fields. Each element represents 16×16 connections made by the 16 subcortical neuronal units or cortical columns per field

intra-column inhibitory connections in W_{ei} and lateral inhibition connections in W_{AC} , respectively (Figure 2b). These three K_i -matrices have default values that reflect the orientation of the current produced by the various types of connections (for details, see Section 4.1 of Hajizadeh et al., 2019). Extending from the approach in our previous study (Hajizadeh et al., 2019), the effect of subject-specific gross anatomy was taken into account in the generation of the MEG signal. To model the effect of the column orientation and distance to the MEG sensor—that is, of the gross anatomy of the cortical folding—the elements of K_i were multiplied by a random number that was specific to each of the 13 cortical fields. This is a simplification, which assumes that the columns in a field have approximately the same orientation and the same distance to the sensor. For notational convenience, these 13 random numbers are represented by a topography matrix T where each row has identical elements and each column is the vector of the 13 field-specific multipliers (see Figure 2c). Further, to model the effect of cross-subject variation of the topography of the cortical surface, multiple T -matrices were generated, so that each T -matrix represented a single subject. Note that in this approach, we are not concerned with describing or reconstructing the effect of the actual cortical gross anatomy of any particular subject but, rather, we are interested in the effect of cross-subject anatomical variation on the ERF response. Thus, each connection has an anatomical multiplier as part of the parameter set A , and this is the product of the elements in T and K_i corresponding to that connection.

The MEG response $R(t)$ of the model is computed as the sum over the synaptic inputs to the excitatory populations, weighted by the Hadamard product of the topography matrix T with the K_i -matrices:

$$R(t) = \sum_{j=1}^N [T \circ K_1 \circ W_{AC}^+ u(t) + T \circ K_2 \circ W_{ei} v(t) + T \circ K_3 \circ W_{AC}^- u(t)]_j. \quad (6)$$

where j runs over the number of cortical columns in the model. The matrices W_{AC}^+ and W_{AC}^- represent the excitatory connections and lateral inhibition of W_{AC} , respectively. The default values for the dynamical and anatomical parameters used in the simulations are listed in Table 1. They were chosen such that the model replicated a typical ERF, with P1m, N1m, and P2m responses. We note that according to our normal-mode characterization (Hajizadeh et al., 2019), each of these responses is fundamentally a property of the whole system and does not have an anatomically localizable generator process. For example, in the model, subtle changes to the internal connections of the parabelt result in significant changes in the activation of the core and belt as they produce the N1m. Thus, the parabelt should be considered to be an integral part of the N1m generator process, even though its direct contribution to this response is small.

2.3 | Simulation experiments

We carried out simulations to test how cross-subject differences in dynamics and/or anatomy of the auditory cortex impacts on the auditory ERF. For this, we randomly varied the parameters of the dynamical equations of the model, denoted by D , as well as the anatomical parameter set A defined in Section 2.2. Further, each set of parameters produced in the randomizations represented an individual subject. The set of dynamical parameters D comprised the weight matrices W_{AC} , W_{ei} , W_{ie} , and W_{ii} . The membrane time constant τ_m is also a dynamical parameter, although it was not varied in the current simulations. For randomizing the D -parameters, the same method was used as in Hajizadeh et al. (2019). That is, for each of the diagonal matrices W_{ei} , W_{ie} , and W_{ii} , a random number was generated from a flat distribution over a predefined range. The elements of the matrix were then multiplied by that number. The matrix W_{AC} , describing the long-range excitatory and lateral inhibitory connections, was generated as a sum of Gaussians with stochastic terms (see Appendix A1 of Hajizadeh et al., 2019). Cross-subject random variation in these column-to-column connection strengths was achieved by regenerating W_{AC} while keeping the Gaussian parameters fixed. Because of the stochastic terms, the overall connectivity pattern remained the same, but weight values varied slightly from subject to subject. To summarize the effect of the randomizations, these essentially altered the balance between excitation and inhibition and they also modified the connectivity patterns at a fine resolution.

The A - and D -parameters were each randomized with five different ranges of the random multiplier. The distribution of random multipliers from each range was evenly distributed around unity, which generated the default value of the modulated parameter (see Table 1). This was achieved by dividing each range into two subsets of random numbers, those larger than unity and those smaller, and then, picking an equal number of multipliers from each subset. For randomizing the D -parameters, the lower bounds of these ranges were chosen as $(1-1/2)$, $(1-1/4)$, $(1-1/8)$, $(1-1/16)$, and unity $(1-0)$. The upper bound was the inverse of the lower bound, and this resulted in the ranges $[0.50, 2.00]$, $[0.75, 1.33]$, $[0.88, 1.14]$, $[0.94, 1.07]$, and $[1.00, 1.00]$ (i.e., no variation). Similarly, for randomizing the A -parameters, we used five different ranges from which the random multipliers were picked to populate the T -matrix (see Section 2.2). These ranges had lower bounds of $(1/8)$, $(1/4)$, $(1/2)$, $(3/4)$, and (1) . Again, the upper bound was the inverse of the lower bound, producing the ranges $[0.12, 8.00]$, $[0.25, 4.00]$, $[0.50, 2.00]$, $[0.75, 1.33]$, and $[1.00, 1.00]$. With five ranges each for the D and A multipliers, there were a total of $5 \times 5 = 25$ combinations of parameter variations. For each combination of A - and D -parameters, 1,000 simulations representing 1,000 subjects were run with parameter values generated randomly with multipliers from the respective ranges of that combination. The resulting ERFs were

TABLE 1 Default dynamical and anatomical parameter values used in the simulations. The matrix W_{AC} includes lateral inhibition (within-field inhibitory) connections and column-to-column excitatory connections. It also contains pre-cortical connections from inferior colliculus (IC) to thalamus, from thalamus to the core areas as well as recurrent connections in IC and thalamus. The cortical intra- and inter-field connections in W_{AC} are defined by Gaussian distributions of the form $Q(x) = r \exp[-(x + \mu + s N(0,1))^2/2\sigma^2]$, with the distance x between column i and j , amplitude r , constant μ , variance σ^2 , stochasticity s of the Gaussian distribution $Q(x)$, and the standard normal distribution $N(0,1)$ (for further details, see Hajizadeh et al., 2019)

Dynamical parameter set D			Value			
τ_m			25 ms			
W_{ei}			1.15			
W_{ie}			1.00			
W_{ii}			0.20			
W_{AC}	(subcortical)	IC recurrent	0.09			
		IC to Thalamus connections	0.015			
		Thalamus recurrent connections	0.09			
		Thalamus to core connections	0.015			
W_{AC}	(cortical)		r	μ	σ^2	s
		Between-field excitatory	0.09	0	1.5	0.2
		Within-field excitatory	0.105	0	2.0	0.4
	Within-field inhibitory	0.09	3.0, -3.0	1.5	0.4	
Anatomical parameter set A			Value			
K_1	feedforward elements		-4			
K_1	feedback elements		20			
K_1	within-field elements		-5			
K_2			2			
K_3			2			

then analyzed in terms of the peak amplitude, peak latency, and 3-dB width of the N1m response. We focused on the N1m, because it is usually the most prominent ERF response generated by the auditory cortex. To account for the time delay due to sub-cortical processing, we added a 35-ms shift to the waveforms. This resulted in the N1m peaking at 100 ms with the default parameter values. Further, going beyond the analysis methods of Hajizadeh et al. (2019), we arithmetically grand-averaged the ERF waveforms across the subjects and inspected the mean waveform as well as its standard deviation.

2.4 | MEG experiments with human subjects

We present MEG data from two separate studies, which were conducted with two different subject populations, with the total number of subjects being 25. Here, we show the analyses of the ERFs from the right hemisphere, and we note that the left-hemispheric ERFs yield the same results with respect to the origin of subject-specificity. Further details on the first study (Experiment A) can be found in Zacharias et al. (2012). Data of the second study (Experiment B) have not been published before, and information on acquisition and pre-analysis is briefly summarized

here. For both studies, subjects were recruited from the academic environment at the Leibniz Institute for Neurobiology and the Otto von Guericke University in Magdeburg. All subjects gave written informed consent to participate in the measurements, and both studies received independent approval by the Ethics Committee of Otto von Guericke University.

Both studies used the same experimental paradigm. Sequences of 1.5-kHz tones with 100-ms duration were presented at a sensation level of 80 dB in separate blocks where each block was characterized by a constant stimulus onset interval (SOI). Two consecutive blocks were separated by a pause of about 30 s, and the order in which the blocks were presented was randomized across subjects. In Experiment A, tones were delivered monaurally to the subjects' left ear, and the SOIs ranged from 0.5 to 10 s in five steps. The recording device was the Magnes 3600 WH system (4-D Neuroimaging) with 248 magnetometers. In Experiment B, tones were delivered binaurally, and 10 different SOIs were used in the 0.25-s to 7-s range. Magnetic fields were measured with the Elekta Neuromag TRIUX system which consists of 102 magnetometers and 204 planar gradiometers (102 measurement locations in total). For the current study, we focused on the measurements performed with the magnetometers of the two systems. The same standard preprocessing procedure (including

artifact rejection, heartbeat correction, filtering and averaging across trials) was applied to the raw data of both experiments (Zacharias et al., 2012). The trial-averaged MEG responses were baseline-corrected (200 ms) and filtered with a band pass of 1–30 Hz (Butterworth, zero phase shift).

As with the simulated responses, the ERFs were analyzed in terms of the peak amplitude, peak latency, and 3-dB width of the N1m response, and also in terms of the grand-averaged ERF and its standard deviation. The measurements relating to the N1m were expressed in terms of histograms, for which we computed 95% confidence intervals for the individual bins. This was achieved via the bootstrap method, that is, random sampling with replacement (Efron, 1979). For this, we constructed 1,000 resamples of the data of the 25 subjects, and then, rejected the 2.5% smallest and the 2.5% largest values in each histogram bin. The remaining values fall in the range depicted in the error bars in the figures.

In Hajizadeh et al. (2019) and in the current study, we ignore the effect of synaptic plasticity, which has a lifetime of several seconds. Hence, for the experimental data, we selected the ERFs from the blocks with the longest SOIs, that is, $SOI = 10$ s in Experiment A and $SOI = 7$ s in Experiment B. The interval between individual stimuli in these blocks was long enough for full recovery from adaptation to occur (see, e.g., Lü et al., 1992; McEvoy et al., 1997; Zacharias et al., 2012). Further, the model offers only an approximation of AC dynamics in terms of damped harmonic oscillators. There are therefore ERF waveforms that it cannot produce but which can nonetheless be observed in a small number of subjects. These are double-peak structures of the N1m waveform and the emergence of a sustained field following the N1m. We excluded such cases from the experimental data to ensure comparability between the simulated waveforms and the experimental data. This led to a rejection of 2/15 subjects in Experiment A and 3/15 subjects in Experiment B.

3 | RESULTS

The current study uses simulations of auditory cortex and experimental data to address how the ERF is shaped by the anatomy of the auditory cortex, on the one hand, and by the system dynamics of the auditory cortex, on the other hand. In simulations, the ERFs of populations of subjects are generated with the assumption that, in each population, the anatomical parameters A and the dynamical parameters D vary across the population in a specified way. We compare simulated ERFs with ERFs from two different MEG experiments (Experiment A and Experiment B; recordings are from magnetometers above the right posterior temporal lobe where the largest ERF response was measured), focusing on the N1m response, the most prominent ERF wave generated in the auditory cortex. We characterize the N1m for each subject

separately in terms of the peak amplitude, the peak latency, and the width of the N1m. Further, we consider the arithmetic mean and arithmetic standard deviation of the waveforms of populations of subjects.

3.1 | Experimental and simulated ERFs

Figure 3a shows trial-averaged MEG responses from individual subjects (thin gray curves) collected in Experiment A and B. Also shown is the corresponding arithmetic mean (grand mean, thick black curve) and arithmetic standard deviation (thick red curve). The subject-specificity of the waveform is apparent in the different peak amplitudes, peak latencies, and waveform widths of the N1m. We note that the peak latency of the N1m in the grand mean occurs at ~ 105 ms, and that the peaks of the individual waveforms are scattered around that value by ± 20 ms, as indicated in the inset of Figure 3a. The standard deviation is time-dependent and shows a pronounced maximum ~ 15 ms before the peak of the N1m of the grand mean. This heteroscedasticity is due the fact that the individual waveforms, and their N1m peaks in particular, predominantly differ by a multiplicative factor, rather than by an additive amount (König et al., 2015; Matysiak et al., 2013).

We simulated cross-subject variability of the ERF by varying the anatomical A -parameters and the dynamical D -parameters of the model (see Table 1). Changing only the anatomical factors while keeping the dynamical parameters constant leads to a large distribution of the peak amplitudes of the P1m, N1m, and P2m responses (Figure 3b). However, the inset shows that there is only a small effect on the peak latencies, which cluster around the respective peak latency in the grand mean waveform. Therefore, the grand mean and the corresponding standard deviation exhibit maxima for the P1m, N1m, and P2m deflections roughly at the same latencies, respectively.

In contrast, a very different pattern is revealed in simulations where the anatomical parameters are fixed but the dynamical parameters are randomized (Figure 3c). In this scenario, the waveforms show very similar P1m deflections. These then evolve into N1m deflections through near-identical positive slopes which fan out into a wide distribution of N1m-peak amplitudes and latencies. There is a strong positive correlation between the peak amplitude and latency of the N1m, as shown in the inset, and the width of the N1m deflection becomes larger with increasing peak amplitude (as in the simulations of Hajizadeh et al., 2019). The spread of the waveforms continues beyond the N1m, and entails a large variety of P2m deflections—contrary to the situation where only anatomical variations are introduced (Figure 3b). The standard deviation shows a pronounced peak at around 130 ms, that is, *after* the N1m peak

of the grand mean. Note that the simulated data in each panel was normalized such that the N1m-peak amplitude of the grand mean of the simulations was equal to the N1m-peak amplitude of the grand mean of the experimental data (350 fT).

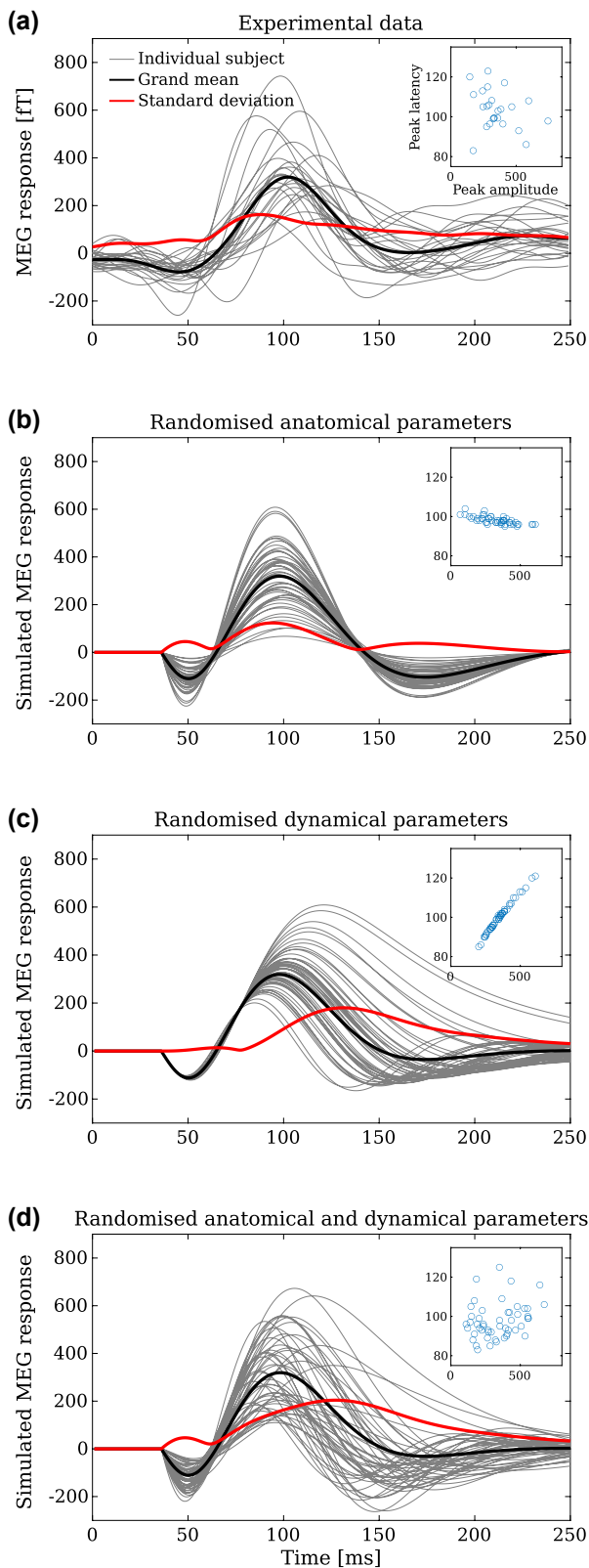


FIGURE 3 Comparison of experimental and simulated waveforms for two illustrative parameter ranges, with the N1m-peak latency plotted against the N1m-peak amplitude in the inset of each panel. (a) Waveforms of individual subjects (thin gray curves; $N = 25$) and their corresponding mean (black curve) and standard deviation (red curve) are shown for the two MEG experiments. The peak amplitude of the N1m varies in a range from about 200 to 800 fT [mean \pm SD: (352 ± 140) fT]. The peak latency ranges from around 80 to 130 ms [(104 ± 10) ms]. (b) Simulated waveforms and the corresponding grand-averaged waveform and standard deviation were generated by randomizing the anatomical parameters A while keeping the dynamical parameters D fixed. A total of 50 waveforms were generated such that the elements of the topography matrix T were randomly picked from the $[0.12, 8.00]$ range. The resulting waveforms are similar in shape. There is large variation in the peak amplitude of the N1m [(327 ± 126) fT] and a narrow variation of the peak latency [(98 ± 2) ms]. (c) The panel shows how varying the dynamical parameters (range of random multiplier $[0.75, 1.33]$) affects the waveform while the anatomical parameters were fixed. There is a strong positive correlation between the amplitude [(352 ± 90) fT] and latency [(100 ± 8) ms] of the N1m peak, as well as between peak amplitude and width of the N1m. (d) The simultaneous randomization of both D - and A -parameters leads to a set of ERF waveforms that resembles that obtained in the experiments and shown in (a). There is a large variation in the N1m-peak amplitude [338 ± 150] fT as well as in the N1m-peak latency [98 ± 9] ms. Note that the simulation results here show a random subset of 50 single-subject responses taken from the 1,000 single-subject simulations, which are the basis of the results shown in the subsequent figures

However, neither of the waveform dispersions in Figure 3b,c bears a close resemblance with that observed in the MEG experiments (Figure 3a). The dispersion due to anatomical variations (Figure 3b) resembles the experimental results in that the standard deviation peaks before the N1m of the grand mean. Nevertheless, the peak latencies of the simulated N1m and P2m are too concentrated. A better correspondence with experimental results is achieved in Figure 3d, which shows simulations where both the anatomical and the dynamical parameters were randomized simultaneously using the respective parameter ranges of the simulations shown in Figure 3b,c. This leads to a wider dispersion of the individual peak latencies of the N1m and P2m, as also demonstrated in the inset, and to a more realistic spread in the N1m widths. These observations suggest that both anatomy *and* dynamics might be the cause of the subject-specificity of ERFs.

3.2 | The N1m response

The dependence of the N1m response on anatomy and dynamics was investigated systematically by using five different ranges of variation for the A -parameters and another five for the D -parameters. Thus, parameter variations were generated in a total of $5 \times 5 = 25$ combinations of parameter

ranges. For each range combination, 1,000 simulated single subjects were generated. The ERF of each subject was characterized in terms of the peak amplitude, the peak latency, and the 3-dB width of the N1m. The normalized distribution of each of these measures was then compared to the normalized distribution of the corresponding measure from the experimental data shown in Figure 3a. The results are shown in Figure 4, where the rows represent the results for the peak amplitude (top), peak latency (middle), and waveform width (bottom) of the N1m, and where the five columns represent the five variation ranges of the *D*-parameters in ascending order from left to right. Within each panel, the five colored curves represent the distributions of the N1m measure, each one gained for a specific range of *A*-parameters.

The gray histograms in the top row of Figure 4 show the distribution of the peak amplitude of the experimentally measured N1m (bin width 100 fT). The distribution is skewed and it has a maximum at the [200–300] fT bin. The corresponding distributions for the simulated ERFs exhibit no *A*-dependence for the largest *D*-range (Figure 4e), as is evident in the colored distribution curves resembling each other. As the *D*-range is decreased (moving from right to left), the amplitude distributions for the different *A*-ranges become increasingly diverse for each *D*-range. The best match between simulations and experiments is achieved with the *A*-ranges with the randomization factor [0.25, 4.00] (purple curves) and [0.12, 8.00] (orange curves). These produce N1m-amplitude distributions which are similar across all the *D*-ranges. On the basis of these results, the subjects have individual cortical anatomies, but no conclusion can be drawn on the presence of dynamical variations across subjects.

The distributions of the peak latency and width of experimentally measured N1m are shown in Figure 4 in the histograms of the middle and bottom row, respectively. The latency distribution is skewed, whereas the width distribution is symmetrical. In both cases, the distributions for the different *A*-ranges closely resemble each other for any given *D*-range, with the one exception of the leftmost panel of the latency row (Figure 4f), where the *D*-range is [1.00, 1.00], that is, when there is no randomization of the dynamical parameters. The best match between experiment and simulation occurs with the *D*-range [0.75, 1.33] both in the case of peak latency (Figure 4i) and in the case of waveform width (Figure 4n). From these findings, we conclude that the subject-specific peak latency and width of the N1m response is explained by cross-subject variations in dynamical parameters, rather than by variations in anatomical factors. This corroborates our observations in Figure 3 that dynamical variations are needed to produce N1m latency variations across subjects.

Figure 5 summarizes the similarity between experimental and simulated data shown in Figure 4. This similarity was quantified separately for each *D*-range and *A*-range combination through the histogram intersection algorithm (Swain &

Ballard, 1991). The similarity results for the peak amplitude, peak latency, and width of the N1m response are shown as similarity maps in Figure 5a–c, respectively. The *x* and *y* axis represent the *D*-ranges and *A*-ranges, respectively, and the color codes the similarity measure, with red elements referring to high and blue elements to low similarity. For the peak amplitude of the N1m (Figure 5a), we note a high similarity across many *D*- and *A*-ranges, with the exception of the narrowest ones (blue panels in the bottom left corner). For the peak latency (Figure 5b), there is a narrow, vertical band of high similarity stretching across all *A*-ranges at the *D*-range of [0.75, 1.33]. Likewise, for the N1m width (Figure 5c), there is a similar vertical band at the *D*-range of [0.75, 1.33]. To identify an overall similarity pattern, the three similarity maps have been averaged in Figure 5d. The dark orange element of this mean map shows which *D*- and *A*-range combination yields the overall best match between experimental and simulated data. This indicates that the subject-specificity of the ERFs is not only based on anatomical variations across subjects, but it also reflects subject-specific dynamics of the auditory cortex.

3.3 | The grand-averaged ERF versus the standard deviation

The correspondence between experimental and simulated data can be examined by looking at the entire ERF waveform rather than at singular time points or deflections such as the N1m response, as was done above. For this broader examination, we used two measures: the ERF waveform grand-averaged across subjects, and the corresponding standard deviation. The results are shown in Figure 6, where the standard deviation is plotted against the grand mean for each of the 25 *D*- and *A*-range combinations (red curves). The five columns represent the five *D*-ranges in increasing order from left (narrow range) to right (wide range), and the five rows represent the *A*-ranges in increasing order from bottom (narrow range) to top (wide range). Each panel also shows the same standard-deviation-versus-mean plot (black curves) for the experimental data extracted from the waveforms shown in Figure 3a. As noted above, the experimental data are heteroscedastic, and this is evident in the characteristic balloon shape of the standard-deviation-versus-mean plot.

In Figure 6, the plots for the simulated data come in a variety of patterns, many of them revealing heteroscedasticity, but only a few of them resemble the experimental data. The panels with the light gray background indicate the best matches in terms of the root mean square analysis between the simulated and the experimental data. They are found in the region with large variations of the *A*-parameters and intermediate variations of the *D*-parameters. These best cases largely overlap with the best matches seen in

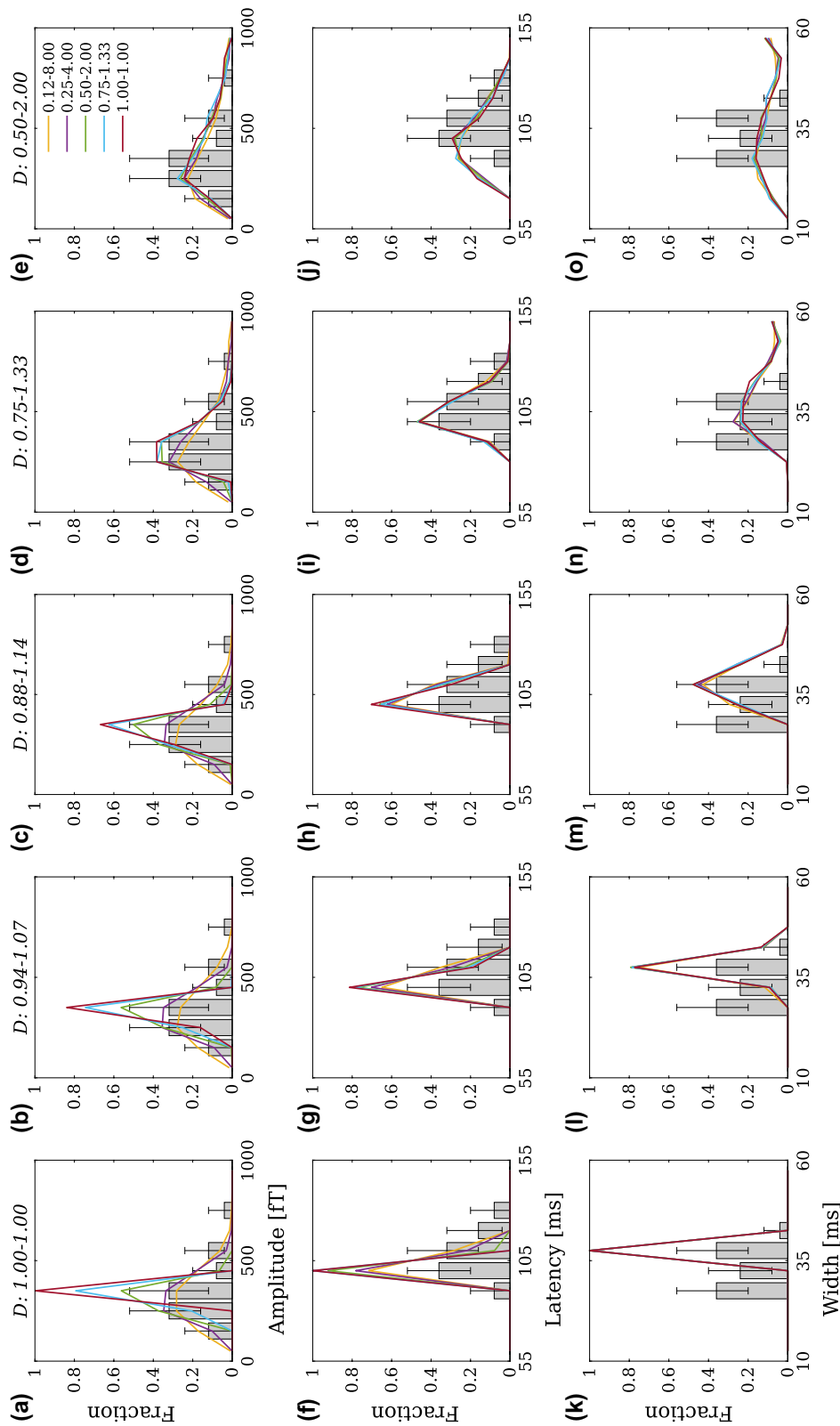


FIGURE 4 Distributions of the peak amplitude (top row), peak latency (middle row), and 3-dB width (bottom row) of the N1m obtained from the experimental data (histograms and simulations (colored lines)). Columns represent the five variation ranges of the dynamical parameters D , in increasing order from left to right. The set of five curves per panel represent five histograms, one for each variation range of the anatomical parameters A (see panel e for color coding of ranges). In each row, the histogram for the experimental results (see Figure 3a) is duplicated across the panels. (a–e) The top row shows the data for the peak amplitude of the N1m. In the simulations, the distribution of the peak amplitude is little affected by varying the A -parameters when the variation of D is large (right-most panels). Conversely, with small D -variations (left-most panels), the distribution has a strong dependence on how A is varied. The best match between simulation and experiment is achieved with the specific A -range of [0.25, 4.00] (purple curves), irrespective of how D is varied. (f–j) The middle row presents the data for the peak latency of the N1m. In simulations, the distributions depend almost exclusively on the variation of the D -parameters, the exception being the case where the D -parameters are kept unchanged (panel f). The best match between the simulation and experimental results are seen when the randomization range for D is [0.75, 1.33] (panel i). (k–o) The bottom row displays the distributions of the N1m widths. As with the peak latency, the widths of the simulated N1m responses depend only on the D -parameters, and the best match with experimental results is obtained with the randomization range [0.75, 1.33] (panel n)

Comparison of simulated and measured N1m responses

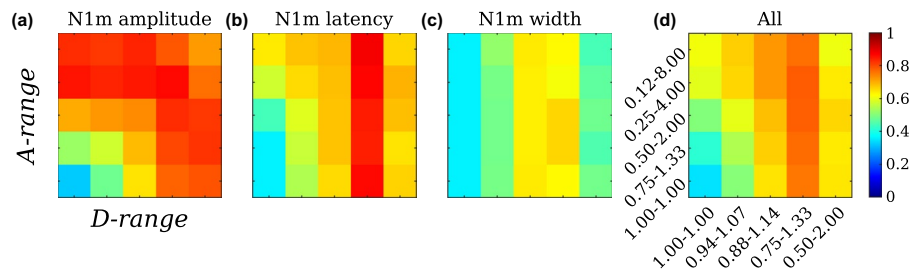


FIGURE 5 Comparison of simulated and measured N1m responses. (a) The distribution of the simulated N1m-peak amplitude resembles the experimental results maximally when the D -parameter variation range is [0.75, 1.33] and the A -parameter range is [0.25, 4.00]. (b) For the peak latency of the N1m, the best correspondence between simulation and experiment is reached with the variation ranges [0.75, 1.33] and [0.12, 8.00] for the D - and A -parameters, respectively. (c) In the case of the N1m width, the best match is reached with the same range [0.75, 1.33] for the D - and A -parameters. (d) Averaging over the previous similarity measures reveals that, overall, the best match between simulation results and experimental data occurs when the variation range of the D -parameters is [0.75, 1.33] and that of the A -parameters is [0.25, 4.00]. Thus, these results are consistent in indicating that the distributions from experimental data are explained by both the cortical anatomy and the cortical dynamics varying from subject to subject. Each panel represents simulations carried out with the 25 combinations of variations of the dynamical parameters D and the anatomical parameters A . The ranges of the D - and A -variations are represented on the abscissa and ordinate, respectively, with the ranges increasing from left to right and from bottom to top. The similarity between distributions was measured with the histogram intersection algorithm (Swain & Ballard, 1991), with high similarity coded in red and low similarity in blue. The actual distributions are shown in Figure 4

Figure 5, where the correspondence between the experimental and simulated N1m response is displayed. All other panels show marked differences between simulations and experiments. For example, simulations displayed in the final column, with the largest D -range [0.50, 2.00], show a wide loop. In these cases, the dynamical parameters dominate the characteristics of the ERFs, leading to a wide distribution of the peak latencies of the N1m, and, consequently, to a shift of the maximum of the standard deviation to a value larger than the maximum of the grand mean (see, e.g., Figure 3d). Similar loops, though less pronounced, and therefore, matching the experimental data, can be seen in the fourth column with a D -range [0.75, 1.33]. When there is no A -variation ([1.00, 1.00], bottom row) and as the range of the D -variation is decreased (from right to left), the relationship between the standard deviation and grand mean approaches homoscedastic behavior, which is finally reached in the case where neither D - nor A -parameters are varied. As the range of the A -variation is increased, homoscedasticity gradually turns into heteroscedasticity, and the relationship between standard deviation and grand mean gradually approaches the pattern identified in the experiments. This pattern indicates both in experiment and simulations, that the individual waveforms in the vicinity of the N1m peak predominantly differ by factors, not by amounts (see Matysiak et al., 2013). In sum, these results confirm the N1m analyses in Section 3.2 according to which the cross-subject variability of the ERF can best be explained by both the cortical anatomy and the dynamical parameters of auditory cortex varying across subjects.

4 | DISCUSSION

While the auditory ERF often comprises a series of landmark deflections identified as the P1m, N1m, and P2m, there is considerable variability across subjects in the peak amplitudes and latencies of these deflections and in the shape of the ERF in general. Indeed, the ERF is much like a fingerprint—in that it is both specific to the individual subject and reproducible across repeated measurements (see Figure 1). We pursued the question of whether this subject-specificity is due to different gross anatomies of the AC, or whether it also reflects subject-specific dynamics. We compared experimental MEG data with simulations of a computational model of the auditory cortex. Our results indicate that the subject-specificity of ERFs is due to a mixture of effects, with both the gross anatomy and dynamics varying across subjects.

4.1 | Main findings

Our model of auditory cortex is based on the anatomical organization of AC in terms of core, belt, and parabelt fields (Hajizadeh et al., 2019; May et al., 2015). There were two sets of parameters that we manipulated in simulations. First, the anatomical parameters represented the effect of the cortical gross anatomy on the generation of the ERF signal, without having an effect on the dynamics. Second, the dynamical parameters were the strengths of the connections between cortical columns. In our simulations, changing these led to a modulation of the balance between excitation and inhibition, and it also affected the patterns of long-range connectivity at

Standard deviation vs. grand mean: Simulated and measured ERFs

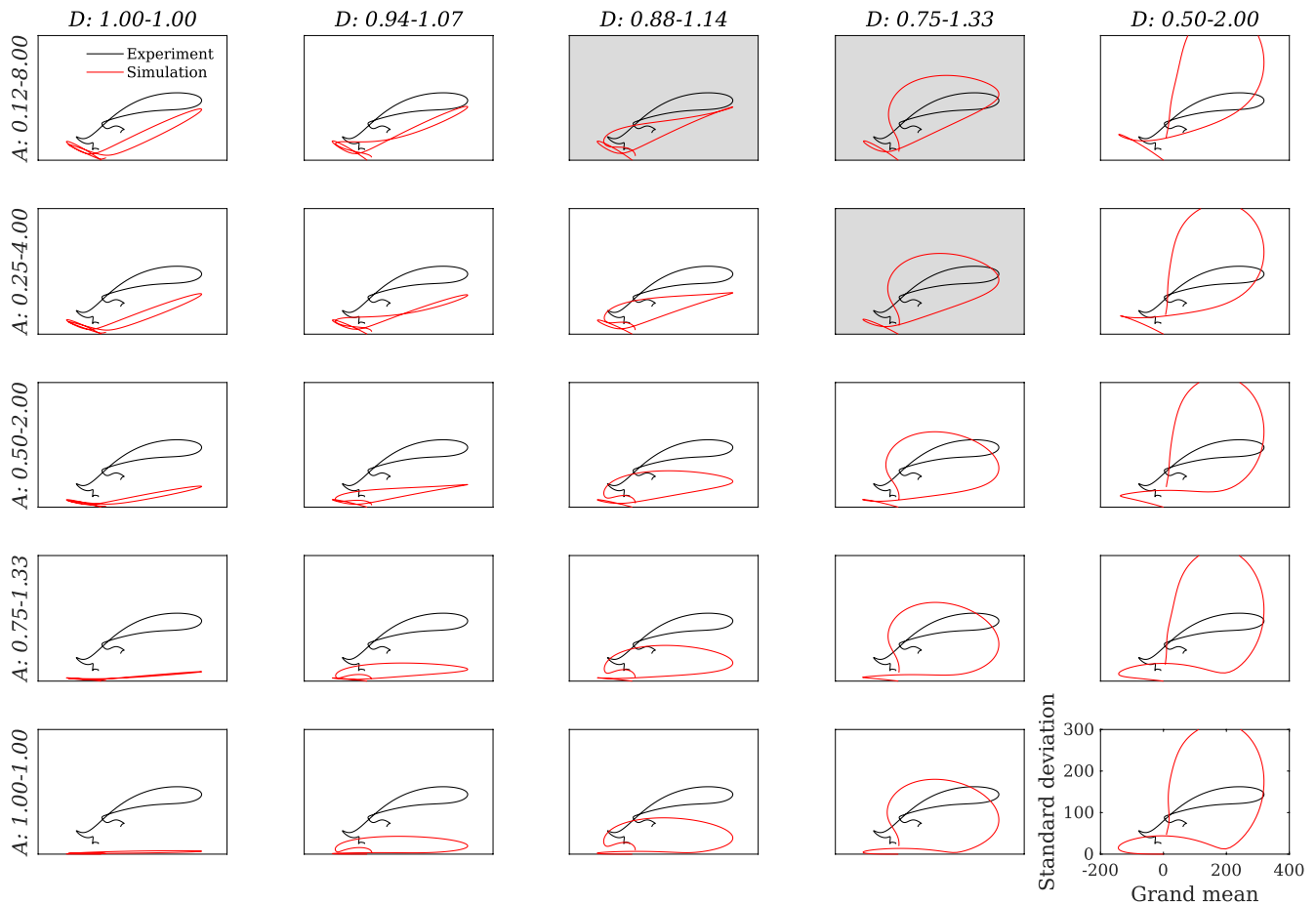


FIGURE 6 Standard deviation versus grand mean: Comparison of simulated and measured ERFs. The experimental data are represented by the black curve, which is replotted in each panel. The simulated data are in red and it represents the 25 combinations of the variation ranges of the dynamical parameters D (abscissa) and anatomical parameters A (ordinate). The width of the D -ranges increases from left to right, and the width of the A -ranges increases from bottom to top. The three best matches between simulations and experiment (based on the root-mean-square differences) are highlighted, and they all display heteroscedastic behavior. These best matches coincide with those shown in Figure 5, and they confirm that the cross-subject variation of the ERF is due to a combination of subject-specific cortical anatomy and subject-specific cortical dynamics. The single-subject ERFs measured in the MEG experiments are shown in Figure 3a. A subset of the waveforms corresponding to the top panel in the fourth column is shown in Figure 3d

a fine resolution. The default parameter values were chosen such that the model replicated a typical ERF. We then introduced variations to both sets of parameter values to generate subject specificity. The simulations thus included 25 sets of ERFs resulting from different combinations of anatomical and dynamical parameter variations. Each set represented the ERFs measured from a subject group, with each individual ERF representing the ERF of a single subject.

The data were analyzed in two ways. First, we considered characteristic quantities of the N1m response, namely its peak amplitude, peak latency, and waveform width. The N1m was chosen for this analysis because it tends to be a landmark response in most subjects, reaching higher peaks than the other ERF responses. On each of these measures,

the experimental data reached a best match with the simulations in which both anatomical and dynamical parameters were varied. Second, we considered the arithmetic mean plotted against the corresponding standard deviation. In simulations, these plots took on a wide variety of shapes, ranging from wide circular orbits to linear-like behavior. Again, we found that the best match between simulations and experiment occurred for the simulations where both the anatomical and dynamical parameters are varied. Furthermore, the variation ranges for the best matches agreed well with each other across the different analyses (compare Figure 5d with Figure 6). In all our comparisons, the experimental results deviated considerably from the predictions at the extremes, whereby either anatomical factors or dynamical factors alone

would explain the subject-specificity of the ERF. Overall, our results agree—unsurprisingly—with the well-established notion that individual subjects have individual cortical gross anatomies. Importantly, they indicate that subjects display cortical dynamics, which is specific to the subject.

We emphasize that our approach relies on analyzing population data and that we are drawing conclusions on subject-specificity on the basis of population data. While this allows us to conclude that both the gross anatomy and the cortical dynamics are subject-specific, it does not, however, permit us to say much about particular individual subjects. For example, our method cannot be used to pinpoint the origin of ERF differences between two subjects, because the number of data points would be too low.

4.2 | N1m-peak latency as an index of group differences in auditory function

On a general level, one might expect that dynamical variation of AC has perceptual consequences. For example, in our model, this variation was brought about by changing the connection strengths between the cortical columns. Connection strengths are fundamental to information processing in that they determine how a stimulus representation is mapped from one stage of processing to the next. Our results suggest that the peak latency of the N1m is a good indicator of cross-subject dynamical variations because it is weakly influenced by anatomy. In contrast, the N1m-peak amplitude is a less sensitive indicator because it is influenced both by dynamical and anatomical variation. Thus, on the group level, it is to be expected that latency should correlate better than amplitude with perception as measured in psychoacoustic experiments, and latency variance should correlate with variance of perception. Indirect evidence suggests that this is indeed the case.

First, the transient sound detection response reported by Mäkinen et al. (2004) and Tiitinen et al. (2005) is an ERF response elicited by long-duration sounds that slowly (over hundreds of milliseconds to seconds) increase in intensity from an imperceptible to a clearly perceptible level. During such looming sound stimulation there emerges a response that is like the N1m in terms of its morphology, polarity, width, and source location in the AC. Furthermore, the timing of the response militates against a fixed-amplitude threshold model, which is also the case with the N1m (Biermann & Heil, 2000). Importantly, the peak latency of this N1m-like response predicts extremely well the behavioral reaction time (RT) indicating sound detection. On the group level, subjects with a short RT display the sound detection response at an earlier latency than subjects with a long RT. Also, the variance in the peak latency of the response is correlated with the variance of the RT.

Second, subjects with musical training tend to produce ERFs that differ from those produced by nonmusicians. Amemiya et al. (2014) presented short melodies to subjects and measured the N1m response elicited by the final tone. They found that musicians have shorter right-hemispheric N1m latencies than nonmusicians. There was no difference in the N1m-peak amplitude between the groups, and the P2m was also unaffected by musical training. The behavioral task was to report on the sense of completeness of the melody. No differences were found between musicians and nonmusicians in this relatively nondemanding task. Similarly, Kuriki et al. (2006) reported that musical training resulted in shorter peak latencies of the N1m, but that the peak latencies of the P1m and P2m were unaffected. Further, Park et al. (2018) found that musicians, compared to nonmusicians, produced N1m responses of shorter peak latency but similar peak amplitude. Interestingly, multiplications of Heschl's gyrus are much more common in musicians (in 90% of the cases, Benner et al., 2017) than in the general population (Marie et al., 2015). Therefore, it seems that both the dynamics and gross anatomy of auditory cortex are changed by musical training.

Further evidence for event-related responses reflecting dynamical variations might be found by looking at the effect of perceptual learning. In this case, the gross anatomy remains unchanged, but cortex undergoes functional changes due to synaptic plasticity (for a review, see Weinberger, 2015). Based on the current results, perceptual learning should be reflected in the group mean of the latency, and/or in the variance of the latency. In line with this, Reincke et al. (2003) found that the N1 and P2 latency shifted earlier and the P2 amplitude became larger when subjects learned to discriminate vowels. This should be contrasted with the results of Tremblay et al. (2001) and Tremblay and Kraus (2002), who found learning to discriminate consonant-vowel syllables was reflected in N1 and P2 amplitudes rather than in their latency. However, the analyses of the above studies were carried out for group means only, and the latency variance was not addressed at the single-subject level.

In sum, previous studies indicate that the latency with which the auditory cortex reaches peak activity in an N1m-like response is a good predictor of sound detection. Further, musical training and perceptual learning affects the aspect of the ERF—the N1m latency—which we suggest is the sensitive indicator of dynamical variation in AC. Also, latency seems to be a consistent indicator of musical training, whereas the N1m amplitude is less so. Thus, the N1m latency appears to be functionally meaningful, reflecting perception and learning. Our results presented in Figures 3–5 explain why this should be the case: variations in cortical dynamics show up in the N1m latency, whereas the amplitude of N1m represents a mixture of dynamical and anatomical effects, which can cancel each other out.

4.3 | On the usefulness of the grand mean of ERF waveforms

The current findings give us pause to consider the usefulness of grand means in ERF research. Grand-averaging the ERF waveform is a continuation of averaging single trial responses. It is rooted in the idea of noise cancelation and in the notion that there is a true population response buried in the noise carried by each single-subject measurement. This practice can be criticized from several viewpoints. First, it almost always relies on arithmetic averaging, which is the convention in ERF research. Arithmetic averaging is based on the assumption of the so-called additive model whereby one response differs from another by a constant. Whereas this assumption is appropriate for single trials (König et al., 2015), the additive model does not apply when computing the grand mean across trial-averaged waveforms. This is because ERFs of individual subjects do not differ from each other by amounts but, rather, they differ predominantly by factors, and follow the so-called mixed model. Therefore, the use of arithmetic averaging is not recommended for comparing the grand mean of ERFs across stimuli or conditions; instead geometric grand-averaging—or, more precisely, arithmetic averaging of the asinh-transformed data—should be used (König et al., 2015; Matysiak et al., 2013). Further, the test–retest reliability of event-related responses of individual subjects is high, meaning that single-trial averaging is successful in removing noise and revealing a response that is stable. These results taken together with the findings of the current study imply that response stability emanates from an unchanging gross anatomy of each subject and from the cortical dynamics of each subject remaining stable. That is, each subject does not introduce noise that should be removed through further cross-subject averaging. If cross-subject variation of ERFs reflected individual gross anatomies only, then, geometric grand-averaging might arguably produce a waveform that represented the true dynamics of the brain, undistorted by measurement noise. However, as the ERF of each subject emerges from the subject-specificity of both gross anatomy and dynamics, it is unclear what the grand mean represents. It blurs the waveforms of individual subjects, and it is unlikely to represent a response produced by average dynamics in an average gross anatomy, even assuming that such notions are useful and can be defined. Therefore, when using grand-averaging, one must ensure that the end justifies the means. We used arithmetic averaging (see Figure 3) to investigate the origin of the subject-specificity of ERFs. If the focus is on differences between groups or experimental conditions, then, geometric averaging or the arithmetic averaging of asinh-transformed waveforms should be employed.

4.4 | Outlook

Our current modeling approach is, to our knowledge, unique for two reasons. First, we are including a description of the entire AC, as opposed to concentrating on a specific cortical field (e.g., A1). As described in our previous study (Hajizadeh et al., 2019), this allows us to characterize the ERF as a holistic waveform being generated by the entire system of the AC, rather than as a series of responses—P1m, N1m, and P2m—with dedicated, spatially constrained generator processes. Second, we are interested in describing subject-specific dynamics and subject-specific anatomical effects on the generation of the auditory ERF, and how these are reflected in group statistics and grand-averaged responses. These strengths should be juxtaposed with the simplifications of the model: For example, it is based on the anatomy of the macaque monkey AC and gross anatomy is modeled as a set of random variables. These simplifications might be seen as springboards for further development, given elasticity by the strengths of our approach. The anatomy of the human auditory cortex is at present unknown in the detail of the macaque monkey (Baumann et al., 2013; Besle et al., 2019; Norman-Haignere et al., 2019). We are not certain how many fields the human AC has and what the connection patterns between the fields are. In contrast, the gross anatomy of subjects is accessible through high-resolution MRI imaging (Moerel et al., 2014). It might be possible to adapt our approach to assist in the mapping of the cortical fields of human AC. This could involve constructing forward models, with multiple MEG sensor locations, on the basis of actual subject-specific cortical topographies. It should be possible to project various field constellations onto these topographies to model the ERF. This might allow one to find the best fits between simulation and experiment in terms of number of fields and general connection patterns between fields, with the constraint that these features are shared across subjects.

Although the current model includes multiple feedback pathways, it is dynamically driven by the afferent input. Looking beyond the auditory cortex, we suspect that the auditory cortex might function as a forum where oscillations driven by the bottom-up sensory input mix with oscillations driven by top-down internal models generated beyond the parabelt. This could be one of the ways that the cognitive auditory cortex (Scheich et al., 2007) manifests itself dynamically and where the top-down-driven oscillations function as filters for the oscillations driven by the afferent input (Morillon & Schroeder, 2015) or vice versa. Thus, the approach of the current study might be extended to examine the auditory ERF—and therefore, the functioning of auditory cortex—in terms of this mixing of these two kinds of oscillations.

ACKNOWLEDGMENT

Open Access funding enabled and organized by ProjektDEAL. WOA Institution: Leibniz-Institut für Neurobiologie Blended DEAL : ProjektDEAL.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

Aida Hajizadeh: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing-original draft; Writing-review & editing. **Artur Matysiak:** Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Writing-review & editing. **André Brechmann:** Conceptualization; Investigation; Methodology; Resources; Writing-review & editing. **Reinhard König:** Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Writing-review & editing. **Patrick J. C. May:** Conceptualization; Funding acquisition; Investigation; Methodology; Project administration; Supervision; Writing-review & editing.

ORCID

Artur Matysiak  <https://orcid.org/0000-0002-1395-8520>

REFERENCES

- Ahonen, L., Huotilainen, M., & Brattico, E. (2016). Within- and between-session replicability of cognitive brain processes: An MEG study with an N-back task. *Physiology & Behavior, 158*, 43–53. <https://doi.org/10.1016/j.physbeh.2016.02.006>
- Amemiya, K., Karino, S., Ishizu, T., Yumoto, M., & Yamasoba, T. (2014). Distinct neural mechanisms of tonal processing between musicians and non-musicians. *Clinical Neurophysiology, 125*, 738–747. <https://doi.org/10.1016/j.clinph.2013.09.027>
- Atcherson, S. R., Gould, H. J., Pousson, M. A., & Prout, T. M. (2006). Long-term stability of N1 sources using low-resolution electromagnetic tomography. *Brain Topography, 19*, 11–20. <https://doi.org/10.1007/s10548-006-0008-8>
- Baumann, S., Petkov, C. I., & Griffiths, T. D. (2013). A unified framework for the organization of the primate auditory cortex. *Frontiers in Systems Neuroscience, 7*, 11. <https://doi.org/10.3389/fnsys.2013.00011>
- Benner, J., Wengenroth, M., Reinhardt, J., Stippich, C., Schneider, P., & Blatow, M. (2017). Prevalence and function of Heschl's gyrus morphotypes in musicians. *Brain Structure & Function, 222*, 3587–3603. <https://doi.org/10.1007/s00429-017-1419-x>
- Besle, J., Mougou, O., Sánchez-Panchuelo, R.-M., Lanting, C., Gowland, P., Bowtell, R., Francis, S., & Krumbholz, K. (2019). Is human auditory cortex organization compatible with the monkey model? Contrary evidence from ultra-high-field functional and structural MRI. *Cerebral Cortex, 29*, 410–428. <https://doi.org/10.1093/cercor/rbhy267>
- Biermann, S., & Heil, P. (2000). Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex. *Journal of Neurophysiology, 84*, 2426–2439. <https://doi.org/10.1152/jn.2000.84.5.2426>
- Brody, D. A., Terry, F. H., & Ideker, R. E. (1973). Eccentric dipole in a spherical medium: Generalized expression for surface potentials. *IEEE Transactions on Biomedical Engineering, 20*, 141–143. <https://doi.org/10.1109/TBME.1973.324177>
- Brosch, M., & Schreiner, C. E. (1997). Time course of forward masking tuning curves in cat primary auditory cortex. *Journal of Neurophysiology, 77*, 923–943. <https://doi.org/10.1152/jn.1997.77.2.923>
- Brosch, M., & Schreiner, C. E. (2000). Sequence sensitivity of neurons in cat primary auditory cortex. *Cerebral Cortex, 10*, 1155–1167. <https://doi.org/10.1093/cercor/10.12.1155>
- Dalebout, S. D., & Robey, R. R. (1997). Comparison of the intersubject and intrasubject variability of exogenous and endogenous auditory evoked potentials. *Journal of the American Academy of Audiology, 8*(5), 342–354.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics, 7*, 1–26. <https://doi.org/10.1214/aos/1176344552>
- Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B. T. T., Mohlberg, H., Amunts, K., & Zilles, K. (2008). Cortical folding patterns and predicting cytoarchitecture. *Cerebral Cortex, 18*, 1973–1980. <https://doi.org/10.1093/cercor/bhm225>
- Hackett, T. A., de la Mothe, L. A., Camalier, C. R., Falchier, A., Lakatos, P., Kajikawa, Y., & Schroeder, C. E. (2014). Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: Refining the hierarchical model. *Frontiers in Neuroscience, 8*, 72. <https://doi.org/10.3389/fnins.2014.00072>
- Hajizadeh, A., Matysiak, A., May, P., & König, R. (2019). Explaining event-related fields by a mechanistic model encapsulating the anatomical structure of auditory cortex. *Biological Cybernetics, 113*, 321–345. <https://doi.org/10.1007/s00422-019-00795-9>
- Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—Theory, instrumentation, and applications to non-invasive studies of the working human brain. *Reviews of Modern Physics, 65*, 413–497. <https://doi.org/10.1103/RevModPhys.65.413>
- Heschl, R. L. (1878). *Über die vordere quere Schläfenwindung des menschlichen Großhirns*, Braumüller.
- Hopfield, J. J., & Tank, D. W. (1986). Computing with neural circuits: A model. *Science, 233*, 625–633. <https://doi.org/10.1126/science.3755256>
- Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 11793–11799. <https://doi.org/10.1073/pnas.97.22.11793>
- König, R., Matysiak, A., Kordecki, W., Sielużycki, C., Zacharias, N., & Heil, P. (2015). Averaging auditory evoked magnetoencephalographic and electroencephalographic responses: A critical discussion. *The European Journal of Neuroscience, 41*, 631–640. <https://doi.org/10.1111/ejn.12833>
- Kuriki, S., Kanda, S., & Hirata, Y. (2006). Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. *The Journal of Neuroscience, 26*, 4046–4053. <https://doi.org/10.1523/JNEUROSCI.3907-05.2006>
- Loebel, A., Nelken, I., & Tsodyks, M. (2007). Processing of sounds by population spikes in a model of primary auditory cortex. *Frontiers in Neuroscience, 1*, 197–209. <https://doi.org/10.3389/neuro.01.1.1.015.2007>

- Lü, Z. L., Williamson, S. J., & Kaufman, L. (1992). Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Research*, *572*, 236–241. [https://doi.org/10.1016/0006-8993\(92\)90475-o](https://doi.org/10.1016/0006-8993(92)90475-o)
- Mäkinen, V., May, P., & Tiitinen, H. (2004). Transient brain responses predict the temporal dynamics of sound detection in humans. *NeuroImage*, *21*, 701–706. <https://doi.org/10.1016/j.neuroimage.2003.10.009>
- Marie, D., Jobard, G., Crivello, F., Perchey, G., Petit, L., Mellet, E., Joliot, M., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2015). Descriptive anatomy of Heschl's gyri in 430 healthy volunteers, including 198 left-handers. *Brain Structure & Function*, *220*, 729–743. <https://doi.org/10.1007/s00429-013-0680-x>
- Matysiak, A., Kordecki, W., Sieluzycycki, C., Zacharias, N., Heil, P., & König, R. (2013). Variance stabilization for computing and comparing grand mean waveforms in MEG and EEG. *Psychophysiology*, *50*, 627–639. <https://doi.org/10.1111/psyp.12047>
- May, P. J., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*, 66–122. <https://doi.org/10.1111/j.1469-8986.2009.00856.x>
- May, P. J., & Tiitinen, H. (2013). Temporal binding of sound emerges out of anatomical structure and synaptic dynamics of auditory cortex. *Frontiers in Computational Neuroscience*, *7*, 152. <https://doi.org/10.3389/fncom.2013.00152>
- May, P. J., Westö, J., & Tiitinen, H. (2015). Computational modelling suggests that temporal integration results from synaptic adaptation in auditory cortex. *The European Journal of Neuroscience*, *41*, 615–630. <https://doi.org/10.1111/ejn.12820>
- McEvoy, L., Levänen, S., & Loveless, N. (1997). Temporal characteristics of auditory sensory memory: Neuromagnetic evidence. *Psychophysiology*, *34*, 308–316. <https://doi.org/10.1111/j.1469-8986.1997.tb02401.x>
- Michalewski, H. J., Prasher, D. K., & Starr, A. (1986). Latency variability and temporal interrelationships of the auditory event-related potentials (N1, P2, N2, and P3) in normal subjects. *Electroencephalography and Clinical Neurophysiology*, *65*, 59–71. [https://doi.org/10.1016/0168-5597\(86\)90037-7](https://doi.org/10.1016/0168-5597(86)90037-7)
- Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience*, *8*, 225. <https://doi.org/10.3389/fnins.2014.00225>
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, *1337*, 26–31. <https://doi.org/10.1111/nyas.12629>
- Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. *NeuroImage*, *13*, 684–701. <https://doi.org/10.1006/nimg.2000.0715>
- Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Bio-Medical Engineering*, *46*, 245–259. <https://doi.org/10.1109/10.748978>
- Näätänen, R. (1992). *Attention and brain function*. Lawrence Erlbaum Associates.
- Norman-Haignere, S. V., Kanwisher, N., McDermott, J. H., & Conway, B. R. (2019). Divergence in the functional organization of human and macaque auditory cortex revealed by fMRI responses to harmonic tones. *Nature Neuroscience*, *22*, 1057–1060. <https://doi.org/10.1038/s41593-019-0410-7>
- Park, J. M., Chung, C. K., Kim, J. S., Lee, K. M., Seol, J., & Yi, S. W. (2018). Musical expectations enhance auditory cortical processing in musicians: A magnetoencephalography study. *Neuroscience*, *369*, 325–335. <https://doi.org/10.1016/j.neuroscience.2017.11.036>
- Rademacher, J., Morosan, P., Schleicher, A., Freund, H. J., & Zilles, K. (2001). Human primary auditory cortex in women and men. *NeuroReport*, *12*, 1561–1565. <https://doi.org/10.1097/00001756-200106130-00010>
- Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*, 781–791. [https://doi.org/10.1016/s0926-6410\(03\)00202-7](https://doi.org/10.1016/s0926-6410(03)00202-7)
- Sandman, C. A., & Patterson, J. V. (2000). The auditory event-related potential is a stable and reliable measure in elderly subjects over a 3-year period. *Clinical Neurophysiology*, *111*, 1427–1437. [https://doi.org/10.1016/s1388-2457\(00\)00320-5](https://doi.org/10.1016/s1388-2457(00)00320-5)
- Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, *32*, 11–22. <https://doi.org/10.1088/0031-9155/32/1/004>
- Scheich, H., Brechmann, A., Brosch, M., Budinger, E., & Ohl, F. W. (2007). The cognitive auditory cortex: Task-specificity of stimulus representations. *Hearing Research*, *229*, 213–224. <https://doi.org/10.1016/j.heares.2007.01.025>
- Segalowitz, S. J., & Barnes, K. L. (1993). The reliability of ERP components in the auditory oddball paradigm. *Psychophysiology*, *30*, 451–459. <https://doi.org/10.1111/j.1469-8986.1993.tb02068.x>
- Shaw, M. E., Hämäläinen, M. S., & Gutschalk, A. (2013). How anatomical asymmetry of human auditory cortex can lead to a rightward bias in auditory evoked fields. *NeuroImage*, *74*, 22–29. <https://doi.org/10.1016/j.neuroimage.2013.02.002>
- Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, *7*, 11–32. <https://doi.org/10.1007/BF00130487>
- Tiitinen, H., Mäkinen, V. T., Kicić, D., & May, P. J. (2005). Averaged and single-trial brain responses in the assessment of human sound detection. *NeuroReport*, *16*, 545–548. <https://doi.org/10.1097/00001756-200504250-00005>
- Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research*, *45*, 564–572. [https://doi.org/10.1044/1092-4388\(2002\)045](https://doi.org/10.1044/1092-4388(2002)045)
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1–P2 complex after speech-sound training. *Ear and Hearing*, *22*, 79–90. <https://doi.org/10.1097/00003446-200104000-00001>
- von Economo, C., & Horn, L. (1930). Über Windungsrelief, Maße und Rindenarchitektonik der Supratemporalfläche, ihre individuellen und ihre Seitenunterschiede. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, *130*, 678–757. <https://doi.org/10.1007/BF02865945>
- Weinberger, N. M. (2015). New perspectives on the auditory cortex: Learning and memory. *Handbook of Clinical Neurology*, *129*, 117–147. <https://doi.org/10.1016/B978-0-444-62630-1.00007-X>
- Westö, J., May, P. J., & Tiitinen, H. (2016). Memory stacking in hierarchical networks. *Neural Computation*, *28*, 327–353. https://doi.org/10.1162/NECO_a_00803

- Williamson, S. J., & Kaufman, L. (1981). Biomagnetism. *Journal of Magnetism and Magnetic Materials*, 22, 129–201. [https://doi.org/10.1016/0304-8853\(81\)90078-0](https://doi.org/10.1016/0304-8853(81)90078-0)
- Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12, 1–24. [https://doi.org/10.1016/S0006-3495\(72\)86068-5](https://doi.org/10.1016/S0006-3495(72)86068-5)
- Zacharias, N., König, R., & Heil, P. (2012). Stimulation-history effects on the M100 revealed by its differential dependence on the stimulus onset interval. *Psychophysiology*, 49, 909–919. <https://doi.org/10.1111/j.1469-8986.2012.01370.x>
- Zhang, Z. (1995). A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres.

Physics in Medicine and Biology, 40, 335–349. <https://doi.org/10.1088/0031-9155/40/3/001>

How to cite this article: Hajizadeh A, Matysiak A, Brechmann A, König R, May PJC. Why do humans have unique auditory event-related fields? Evidence from computational modeling and MEG experiments. *Psychophysiology*. 2021;00:e13769. <https://doi.org/10.1111/psyp.13769>