

European Journal of Neuroscience, Vol. 41, pp. 615-630, 2015

# Computational modelling suggests that temporal integration results from synaptic adaptation in auditory cortex

# Patrick J. C. May, Johan Westö and Hannu Tiitinen

Department of Biomedical Engineering and Computational Science (BECS), School of Science, Aalto University, P.O. Box 12200, FI-00076 Aalto, Finland

Keywords: combination sensitivity, forward facilitation, mismatch response, N1m response, stimulus-specific adaptation

#### Abstract

Incoming sounds are represented in the context of preceding events, and this requires a memory mechanism that integrates information over time. Here, it was demonstrated that response adaptation, the suppression of neural responses due to stimulus repetition, might reflect a computational solution that auditory cortex uses for temporal integration. Adaptation is observed in single-unit measurements as two-tone forward masking effects and as stimulus-specific adaptation (SSA). In non-invasive observations, the amplitude of the auditory N1m response adapts strongly with stimulus repetition, and it is followed by response recovery (the so-called mismatch response) to rare deviant events. The current computational simulations described the serial core-belt-parabelt structure of auditory cortex, and included synaptic adaptation is sufficient for columns to respond selectively to tone pairs and complex tone sequences. These responses were defined as combination sensitive, thus reflecting temporal integration, when a strong response to a stimulus sequence was coupled with weaker responses both to the time-reversed sequence and to the isolated sequence elements. The temporal complexity of the stimulus seguest that while synaptic adaptation produces facilitation and suppression effects, including SSA and the modulation of the N1m response, its functional significance may actually be in its contribution to temporal integration. This integration seems to benefit from the serial structure of auditory cortex.

# Introduction

Because most auditory objects have a temporal as well as a spectral structure, the auditory system is required to form representations of these objects by integrating memory representations of previous stimulus events with incoming stimulation. While the underlying neural mechanisms of this process are unknown, the end result manifests itself on the perceptual level as integration and segregation phenomena in auditory scene analysis (ASA; Bregman, 1990). A necessary aspect of neural responses pointing to temporal integration is that they depend on the historical context of incoming stimulation. Such memory effects can be seen non-invasively in the auditory N1 response measured in electroencephalography (EEG) and in its magnetic counterpart N1m, both peaking some 100 ms after stimulus onset. The amplitude of the N1(m) is inversely related to the interstimulus interval (Lu et al., 1992) and can be markedly diminished already by a single stimulus repetition (Budd et al., 1998; for a review, see May & Tiitinen, 2010). In the oddball stimulation paradigm where frequently occurring 'standard' stimuli are followed by a rare 'deviant' stimulus, the deviant elicits an N1(m) that is

Received 11 September 2014, revised 2 December 2014, accepted 4 December 2014

enhanced in amplitude compared with that elicited by the standards. This enhancement, sometimes called the mismatch negativity (MMN), is usually interpreted as a change detection response indexing auditory sensory memory (Näätänen, 1992; Picton *et al.*, 2000).

A similar pattern can be seen in single-cell recordings in primary auditory cortex. These show forward masking and stimulus-specific adaptation (SSA), whereby the repetition of a stimulus leads to diminished responses and to recovery of the response strength if the repeated stimulus is followed by a stimulus of a different frequency (Calford & Semple, 1995; Brosch & Schreiner, 1997, 2000; Ulanovsky et al., 2003, 2004). Adaptation associated with forward masking and SSA occurs in all stages of the auditory pathway. In the early stages, masking has a fast recovery time constant of tens of milliseconds (Bleeck et al., 2006), and the presence of SSA in non-lemniscal midbrain structures has recently been established (for a review, see Pérez-González & Malmierca, 2014). However, because SSA is weak in the lemniscal pathway leading to auditory cortex, it seems that adaptation associated with forward masking and SSA in cortex is mainly cortical in origin rather than being an effect that is produced subcortically and merely passed on to the response patterns of cortical neurons (Taaseh et al., 2011; Nelken, 2014).

The above evidence points to context sensitivity: response strength seems to reflect the statistical likelihood of the stimulus in

*Correspondence*: Dr P. J. C. May, as above. E-mail: patrick.may@aalto.fi

a time window stretching several seconds into the past (see also Campbell *et al.*, 2010; Klampfl *et al.*, 2012). However, this does not yet show that temporal integration is taking place that would allow for responses selective to the temporal structure of sound. Such evidence is constituted by cells in auditory cortex showing selectivity to tones (McKenna *et al.*, 1989; Brosch & Schreiner, 1997, 2000; Brosch *et al.*, 1999; Brosch & Scheich, 2008; Sadagopan & Wang, 2009) and species-specific vocalizations (Lewicki & Konishi, 1995; Wang *et al.*, 1995; Rauschecker, 1997; Recanzone, 2008) presented in a particular sequence. This points to a mechanism that enables cells to respond *selectively* to the temporal order in which spectral information is delivered to auditory cortex.

Given the dependence on historical context of auditory responses on the one hand and the temporal integration demonstrated by single-cell responses in auditory cortex on the other, do these phenomena have a common denominator? That is, does the memory system indexed by non-invasive responses originating in auditory cortex serve temporal integration necessary for the auditory system to make sense of the sound environment? Our previous work suggested that short-term synaptic adaptation (i.e. activity-dependent depression) might underlie the adaptation of the N1(m) as well as the mismatch response (May et al., 1999; May & Tiitinen, 2007, 2010). More recently, a computational model of auditory cortex was used to show that adaptation might be crucial for the ability of auditory cortex to respond selectively to speech sounds (May & Tiitinen, 2013). This synaptic plasticity explanation of the emergence of selectivity to temporal structure was also outlined in the computational work of Buonomano and colleagues (Buonomano & Merzenich, 1995; Buonomano & Maass, 2009) and has recently been developed to explain how primary auditory cortex responds selectively to paired tones (Goudar & Buonomano, 2014). Here, using our previously described computational model (May & Tiitinen, 2013), we fan out from earlier approaches to explore whether synaptic adaptation could both underlie the temporal integration of tone pairs and longer tone sequences, and be reflected in memory and masking effects found in vivo and non-invasively, in the N1(m) and mismatch responses. Also, we look at how responses indicating temporal integration are distributed over primary (core) and secondary (belt and parabelt) areas of auditory cortex.

#### Materials and methods

#### Model dynamics

Auditory cortex was simulated with a model used in May & Tiitinen (2013), where a detailed description can be found. The dynamical unit of the model was the cortical microcolumn (N = 208), which contained a population of excitatory (pyramidal) cells and a population of inhibitory interneurons. The pooled activity of these populations was described using the Wilson and Cowan firing rate model (Wilson & Cowan, 1972). For each population, the firing rate g depended on the state variable u through a non-linear monotonically increasing function  $g(u) = \tanh(2/3)(u-\theta)$  when  $u > \theta$ , g(u) = 0 otherwise ( $\theta = 0.1$  is a threshold constant). Thus, firing rate was limited to the [0, 1] range. With the vectors  $\mathbf{u} = [u_1 \dots u_N]$  and  $\mathbf{v} = [v_1 \dots v_N]$  denoting the state variables of the excitatory and inhibitory cell populations, respectively, the dynamic equations describing the neural interactions are:

$$\begin{cases} \tau_{\rm m} \dot{\mathbf{u}}(t) = -\mathbf{u}(t) + W_{\rm ee} \cdot g[\mathbf{u}(t)] - W_{\rm ei} \cdot g[\mathbf{v}_{\rm i}(t)] + \mathbf{I}_{\rm aff}(t), \\ \tau_{\rm m} \dot{\mathbf{v}}(t) = -\mathbf{v}(t) + W_{\rm ie} \cdot g[\mathbf{u}(t)] \end{cases}$$
(1)

where  $\tau_m = 30$  ms is the membrane time constant (see Koch *et al.*, 1996),  $W_{\rm ee} > 0$  is the matrix of excitatory connections between the pyramidal populations,  $W_{ie} > 0$  describes the weights from pyramidal to interneuron populations, and  $W_{ei} > 0$  denotes the weights from the interneurons to the pyramidal cell population. The vector  $\mathbf{I}_{\mathrm{aff}}$  describes afferent input arriving from the auditory pathway. Specifically, the spectral analysis carried out by the subcortical auditory pathway (Young, 2008) was modelled by transforming the auditory stimuli into 16-channel spectrograms with a 1-ms time resolution. The channels spanned the frequencies 100-15 557 Hz with logarithmic spacing  $f_{i+1} = 1.4 f_i$  and carried values in the [0, 1] range. This resulted in a crude, tonotopic representation of the time-evolution of the spectral content of the sounds. Thus, for each area of the model receiving (tonotopically organized) input, the frequency channels of the spectrogram were mapped onto sequential elements of  $I_{aff}$ , and for columns in areas not receiving afferent input, the corresponding elements of  $I_{aff}$  were zero.

Synaptic depression is a likely candidate for the mechanism of adaptation (Wehr & Zador, 2003, 2005), although other mechanisms such as potassium currents probably contribute also (Abolafia *et al.*, 2011). The role of synaptic depression is supported by the results of Ulanovsky *et al.* (2004), who found that SSA operates on multiple concurrent time scales, ranging from a few hundred milliseconds to several seconds; these match the several coexisting time constants that describe the recovery from synaptic depression of corticocortical synapses (Tsodyks & Markram, 1997; Varela *et al.*, 1997; Markram *et al.*, 1998). Thus, synaptic depression in the model affected the synapses between the pyramidal cell populations. Specifically, this was realized by modifying  $W_{ee}$  by a time-dependent adaptation term a(t). The effective synaptic weight between columns *i* and *j* was the product  $a_{ij}(t)w_{ij}$  and depended on the presynaptic activity through:

$$\dot{a}_{ij}(t) = \frac{1 - a_{ij}(t)}{\tau_{\rm a}} - k a_{ij}(t) g[u_j(t)], \qquad (2)$$

where  $\tau_a = 1.6$  s is the time constant of adaptation and k = 20 is a constant. This resulted in a relatively fast onset of adaptation (within 100 ms) and a slower recovery time of several seconds (May & Tiitinen, 2013). The current model, which coarsely lumps all cortical layers into one, did not include plasticity of thalamocortical synapses, which show both depression (in layer 4) and facilitation (in layer 2/3; for a review, see Reyes, 2011).

The magnetoencephalographic (MEG) response of the model was assumed to be proportional to the sum of the excitatory inputs to the pyramidal cells, weighted by the synaptic weights and the suppression terms (Okada *et al.*, 1997; May & Tiitinen, 2010).

$$R(t) = \sum_{ij} k_{ij} w_{ij} a_{ij}(t) g[u_j(t)], \qquad (3)$$

where the term  $k_{ij} = 1$  for area-to-area feedforward connections of  $W_{ee}$  and  $k_{ij} = -1$  for feedback connections. This was implemented due to evidence showing that feedback activity contributes to responses that have a polarity opposite to that of the N1m (Garrido *et al.*, 2007).

#### Model structure

As shown in Fig. 1A, the 208 columns of the model were divided into 13 cortical areas ( $N_{\rm F} = 16$  columns per area) that were organized into a core-belt-parabelt structure (Hackett *et al.*, 1998; Kaas



FIG. 1. Connectivity of the model. (A) The schematic diagram shows how the model consisted of three core areas, eight belt areas and two parabelt areas. These interacted via denser and sparser interconnections (denoted by large and small arrowheads, respectively). With afferent input targeting the core areas only, feedforward activation progresses along multiple core-beltparabelt streams. Connections between areas were topographic and bi-directional. (B) The above schematic diagram translated into the  $208 \times 208$ weight matrices  $W_{ee}$  (blue dots) and  $W_{ie}$  (red dots) mediating column-to-column excitatory and functionally inhibitory connections, respectively. Connections between excitatory populations occurred within the column (diagonal values of  $W_{ee}$ ), within each area (16 × 16 diagonal subdivisions of  $W_{ee}$ ) and between areas (off-diagonal subdivisions of  $W_{ee}$ ). Inhibition was local also in the sense that the interneuron population received excitatory input from within the same area only, and hence  $W_{ie}$  had a diagonal structure. Connections below and above the diagonal subdivisions are feedforward and feedback connections, respectively. Dots represent non-zero values.

& Hackett, 2000). As described above, each column constituted one excitatory and one inhibitory population. The connectivity of the model was described on three levels of resolution: connections within a column; those within an area; and those between areas. Synaptic weights were assumed to be strongest within a column, so that the diagonal values of  $W_{ee}$  were set to  $w_{jj} = 6$  (and off-diagonal, lateral values were of magnitude 0.5; see below). With  $I_{aff} \le 1$ ,

this provided for an amplifying effect of afferent input through recurrent excitation, as suggested by the results of Douglas *et al.* (1995). Inhibition was local in the sense that the interneurons of each column projected only to the pyramidal cells of that column (i.e. the only non-zero values of  $W_{ei}$  were on the diagonal,  $w_{jj} = 3.5$ ). The diagonal values of  $W_{ie}$ , that is, the local, interneuron-targeting excitatory connections within the column, were set to a magnitude of 3.5.

Within each of the 16 areas, inter-column connections originated from the pyramidal cell population and thus were described by the 13 subdivisions ( $16 \times 16$  'intra-area' matrices) along the diagonals of  $W_{ee}$  and  $W_{ie}$  (Fig. 1B). The probability of a symmetric, intra-area connection between two columns was  $P_0 = 0.4$ , and had a Gaussian drop-off from the diagonal with a standard deviation of  $\sigma = 0.6 P_0 N_F$  (Levy & Reyes, 2012). These excitatory connections targeted either the pyramidal or interneuron population of the receiving columns, respectively, and therefore were either functionally excitatory (via  $W_{ee}$ ) or functionally inhibitory (via  $W_{ie}$ ). Thus, in this model, 'functionally inhibitory' connections are the excitatory connections made from one column to the inhibitory interneuron population of another column. Such lateral inhibitory interactions are found in A1 (Kurt et al., 2008; Moeller et al., 2010), and were assumed to hold for belt and parabelt areas also. The lateral intra-area elements of Wee and Wie were set to magnitudes 0.5 and 10, respectively. The probability that an intra-area connection was functionally inhibitory (rather than excitatory) was  $P_{\rm inh} = 0.8.$ 

Inter-area connectivity, expressed through the off-diagonal values of  $W_{ee}$ , was based on the results from primates (Hackett *et al.*, 1998; Kaas & Hackett, 2000). Afferent input  $I_{aff}$  targeted three 'core' areas. These were interconnected with each other and with eight 'belt' areas; belt areas, in turn, were interconnected with two 'parabelt' areas (Fig. 1A). Strong and weak connectivity between areas (Hackett et al., 1998) was realized through column-to-column connection probabilities  $P_1 = 0.1$  and  $P_2 = 0.05$ , respectively. As the core areas made no direct connections with the parabelt (Fig. 1A), the result was the presence of multiple core-belt-parabelt streams of connections with a roughly 'rostral' and 'caudal' subdivision. Connections between areas were topographic (de la Mothe et al., 2006) and therefore each inter-area subdivision of  $W_{ee}$ was characterized by a diagonal structure (Gaussian drop-off,  $\sigma = 0.6 PN_{\rm F}$ ). Additional simulations showed that with values of  $P_1$ or  $P_2$  above 0.3, transient stimulation resulted in the model settling into a stable, sustained activity mode after stimulation, that is, an attractor state where  $u_i > 0$ ,  $\forall i$ . Only with  $P_1 < 0.3$  and  $P_2 < 0.3$ did the model return to its baseline state  $u_i = 0$ .

The above parameter values ensured, first, that preferred pure tone stimulation of core region columns elicited a transient response followed by sustained activity of stimulus duration (which is in agreement with the results of Wang et al., 2005; see fig. 1 of May & Tiitinen, 2013). Second, the MEG response to pure tone stimulation peaked at about 100 ms after stimulus onset, and therefore resembled the N1m response. This second point lead us to use stronger inhibition (in terms of the diagonal values of  $W_{ie}$  and  $W_{ei}$ ) than in the May & Tiitinen (2013) study. However, the exact parameter values used here were not necessary for temporal integration to take place. This was verified by a set of simulations where the weight values of  $W_{ee}$ ,  $W_{ie}$  and  $W_{ei}$  were varied in five steps in a 25-200% range in relation to their default values. With separate variations for the diagonal and off-diagonal elements, and by varying the diagonal values of Wie and Wei in tandem, this resulted in  $5^4 = 625$  weight combinations. In a separate set of simulations, the probabilities  $P_{\text{inh}}$  and  $P_0$  were each varied in the [0, 1] range in 0.1-steps, resulting in 121 probability combinations.

#### Stimuli

Stimulus sets comprised pure tones, two-tone combinations and sequences of multiple tones. The stimuli were normalized with respect to their root-mean-square values. As described above, the stimuli were transformed into 16-channel spectrograms, which were then mapped onto the 16 columns of each core area. All tones were of 50-ms duration, had a linear onset and offset ramps of 5 ms, and their frequencies coincided with the frequency channels of the model. The stimulus spectrograms were normalized to unity and presented to the three core areas, with each frequency channel targeting one column per core area.

Tone pairs were presented in isolation and in the oddball stimulation paradigm. In the experiments looking at responses to isolated tone pairs, the 1054-Hz and 1476-Hz tones (representing one frequency step in terms of frequency channels) were used to construct an ascending pair and the corresponding descending pair. The intrapair stimulus-onset asynchrony (SOA) was varied in 12 logarithmic steps in the 50-2360 ms range. In the oddball experiments (schematically illustrated in Fig. 2A), the above ascending pair was used as the standard (P = 0.9) and the descending pair was the deviant (P = 0.1). Two intra-pair SOAs, 600 ms and 50 ms, were used in separate experiments. In both cases, the onset-to-onset inter-pair interval (IPI) between tone pairs was 2000 ms. In a further experiment, demonstrating the 'primitive intelligence' mismatch response, stimuli comprised the ascending tone pairs 753-1054 Hz, 1054-1476 Hz, 1476-2066 Hz and 2066-2893 Hz, as well as the corresponding descending pairs (SOA = 50 ms, IPI = 550 ms). The set of ascending tone pairs was presented as the standard (P = 0.9) so that each pair had a presentation probability of P = 0.225. The descending tone pairs were the deviants (P = 0.1), with P = 0.025for each pair. A modified version of this set-up was also used in which both the standard and deviant comprised ascending as well as descending tone pairs (standards: 753-1054 Hz, 1476-2066 Hz, 1476-1054 Hz, 2893-2066 Hz; deviants in reverse). In all the oddball experiments, the total number of stimulus presentations was 200.

Two experiments used complex sequences of either four or five tones. In each case, a unique choice of four 50-ms tones was first made from a larger set comprising tones with frequencies 536, 753, 1054, 1476, 2066, 2893, 4050 and 5669 Hz. The number of unique choices and therefore the total number of four-tone sequences was 70. Here, we denote this set of unique sequences as  $S_{70}$ . In the first experiment (Fig. 2B), 70 stimulus sets were used where each set contained four combinations: (1) one of the four-tone sequences picked from  $S_{70}$ ; (2) its reversed version; (3) the initial two-tone part of the sequence; and (4) the final two-tone part (inter-tone SOA = 400 ms; sequence length 1250 ms). In the second experiment, S<sub>70</sub> was again used to construct a total of 70 stimulus sets, as shown in Fig. 2C. This time, an individual stimulus set comprised 25 stimuli: the 1054-Hz tone presented in isolation and a total of 24 five-tone sequences. The first four tones of these sequences represented the 24 permutations of the original four-tone sequence picked from  $S_{70}$ , and these were always followed by the 1054-Hz tone (inter-tone SOA = 150 ms; sequence length 650 ms). Thus, the stimuli in each set had the same spectral composition, they all ended with the same tone, and they all differed from each other with respect to the temporal order in which the spectral content was



FIG. 2. Schematic diagrams of the experimental set-ups. (A) In the oddball paradigm, an ascending tone pair (blue) was used as the frequently occurring standard stimulus, and the corresponding descending tone pair (red) was the rarely occurring deviant. The intra-pair stimulus-onset asynchrony (SOA) and the onset-to-onset inter-pair interval (IPI) are shown. (B). For the first experiment utilizing complex sequences, a set of 70 distinct four-tone sequences ( $S_{70}$ ) was constructed using tones of eight different frequencies. Each sequence was accompanied by its reversed version as well as the first and second parts of the sequences were used. The starting point for constructing each set was an individual sequence from  $S_{70}$ . This four-tone sequence provided 24 permutations. The final five-tone sequences were prepared by adding the 1064-Hz tone to each permutation. The final, 25th member of the set was the 1064-Hz tone presented in isolation.

delivered. Two of such sequences were used in an additional oddball experiment looking at the MEG responses elicited by these complex stimuli.

### Analysis

The analysis of the model examined the firing rates of the pyramidal population of each column:  $f(t) = g[u_i(t)]$ ,  $i = 1, ..., 208, f \in [0, 1]$ . For each stimulus and column, the maximum firing rate  $f_{\text{max}}$  was identified in a time window extending from stimulus onset to 350 ms after stimulus ending. The role of synaptic depression was evaluated by contrasting the results gained by using the default model  $\tau_a = 1.6$  s (denoted as 'slowly-decaying adaptation') with the

results from a model where depression was assumed to decay with a fast time constant of  $\tau_a = 100$  ms ('fast-decaying adaptation'). These values fall within the experimentally measured recovery time constants of depression in corticocortical synapses (i.e. tens of milliseconds to several seconds; Tsodyks & Markram, 1997; Varela et al., 1997; Markram et al., 1998).

SSA on the column level was investigated via the oddball paradigm (Ulanovsky et al., 2003). First, the responses to the standards (180 presentations) and deviants (20 presentations) were averaged separately. Second, for each column,  $f_{max}$  was determined from the averaged response to the standard, from the averaged response to the deviant, and from the response elicited by the standard when presented in isolation. Columns were chosen for further analysis if the isolated standard produced a maximum response of  $f_{\text{max}} > 0.1$ . This use of a threshold relates to experimental studies in which SSA is determined using tone frequencies that produce a robust response in the cells being studied (i.e. frequencies close to the best frequency of the cell). The choice of threshold affected the absolute number of SSA columns found, but had little effect on the relative distribution of SSA columns in the different regions of the model. The response to the standard was judged to represent SSA if it satisfied two conditions: (1)  $f_{\text{max}}$  to the standard was smaller than  $f_{\text{max}}$  to the isolated standard (adaptation); and (2)  $f_{\text{max}}$  to the deviant was larger than  $f_{\text{max}}$  to the standard (recovery).

Temporal integration in a column was demonstrated when the column exhibited temporal combination sensitivity (CS). To meet the requirements of CS, a column had to respond to the stimulus as a temporal whole ( $f_{\text{max}} > 0.1$ , i.e. 10% of potential maximum firing rate) in the sense that it produced a much weaker response to the time-reversed version of the stimulus (Wang et al., 1995), and to the first and the second half of the stimulus presented in isolation (Rauschecker, 1997). Accordingly, in the experiments using tone pairs, the responses to the ascending tone pair (1052-1476 Hz) and its descending counterpart were compared with a control set of three other responses: that elicited by the reversed tone pair and those elicited by the tones presented in isolation. A column was considered to exhibit CS and to be temporally sensitive if  $f_{max}$  to the tone pair was more than double that to any of the control set. This threshold criterion is the same as that used in our previous study (May & Tiitinen, 2013), and is in line with the Preference Index calculation introduced by Rauschecker et al. (1995). The proportion of columns meeting the requirement of CS was denoted  $P_{\rm CS}$ . We also calculated Prev, the proportion of columns responding 100% stronger to the forward tone pair than to the reversed version, and  $P_{cnt}$ , the proportion of columns responding 100% stronger to the forward tone pair than to either of the tones presented in isolation. Changing the above  $f_{\text{max}} > 0.1$  criterion for choosing columns for CS analysis affected the absolute values of the proportions P, but left their relative magnitude order intact (i.e.  $P_{\rm rev}$  was always the largest,  $P_{\rm CS}$ was always approximately equal to  $P_{cnt}$ ).

Experiments were conducted to compare the behaviour of the model with results from primate experiments demonstrating another indicator of temporal integration, namely that of forward facilitation (Brosch et al., 1999; Brosch & Scheich 2008). These experiments use two-tone combinations and examine how the response to the second tone is affected by parametrically changing the first tone. In the current study, the SOAs of the ascending and descending tone pairs were varied in 12 steps on a logarithmic scale from 50 to 2360 ms. Columns were chosen for further analysis if the second tone presented in isolation produced a response ( $f_{\text{max}} > 0.1$ ), and thus was in the receptive field of the column. Using this maximal response as a baseline  $f_{\text{base}}$ , a column was considered to exhibit

Temporal integration 619

response facilitation if the presentation of the first tone at any SOA resulted in an  $f_{\text{max}}$  to the second tone such that  $f_{\text{max}} > 1.1*f_{\text{base}}$ . For each such column, we obtained the magnitude of the facilitation  $f_{\text{max}}/f_{\text{base}}$ , the SOA that produced the largest facilitation, and the longest facilitating SOA.

The analyses of the responses to the complex tone sequences mirrored those used to examine the two-tone responses. For the fourtone sequences (Fig. 2B), the response in terms of  $f_{\text{max}}$  to each sequence was compared with that elicited by the reversed sequence as well as to those evoked by the first and second half of the sequence. Using the same criteria for CS responses as used in the tone-pair case, the results were quantified as  $P_{\rm CS}$ , the proportion of columns exhibiting CS. In the case of the five-tone sequences (Fig. 2C), we adapted the measure used to quantify forward facilitation in the tone-pair case. Thus, the response to the final 1054-Hz tone of a sequence was compared with the 1054-Hz tone presented in isolation as well as to the responses elicited by the other 23 unique tone combinations. Temporal integration was demonstrated if a column responded selectively to only one of the sequences, that is,  $f_{\text{max}}$  was more than double the corresponding values produced by the isolated tone and the other sequences.

The MEG responses generated by the model were analysed in a variety of experiments employing the oddball paradigm. In these, the responses to the standards (N = 180) and deviants (N = 20) were averaged separately. These responses exhibited an N1m-like deflection in that they peaked at about 100 ms and showed a strong amplitude dependence on stimulus interval. The peak of this response was quantified in terms of amplitude and latency. A more faithful replication of experimentally observed event-related responses, including the P50m and P200m deflections, is beyond the scope of the current study. This would require modelling the contours of auditory cortex and the targeting zones of synaptic inputs on dendritic trees within the cortical column, all of which affect the polarity of the MEG signal (Hämäläinen et al. 1993; Okada et al., 1997; May & Tiitinen, 2010). To link the current results with previous experimental efforts concentrating on the mismatch response, the response to the deviant was subtracted from that elicited by the standard. Further, the columns generating the standard and deviant responses were identified separately by pinpointing the columns that were active  $(f_{\text{max}} > 0)$ during respective standard-response and deviant-response time windows. In each case, the window was defined as the period during which the MEG exceeded a limit value calculated as 10% of the mean peak amplitude of the standard and deviant response. In the case of the response to the deviant, this window included the mismatch response in all cases. It was determined which set of columns C were activated  $(f_{\text{max}} > 0)$  by any one of 16 isolated tones in the 100-15 557 Hz hearing range of the model, where the frequencies of the tones coincided with the frequency channels of the model. The set C also included those columns that were activated by a 50-ms burst of white noise or by a 50-ms complex tone including the frequencies used in constructing the standard and deviant stimuli.

#### Results

#### Basic response patterns to isolated and repeated tones

To demonstrate the basic response pattern of the model, a brief 50ms pure tone of frequency 1054 Hz was used as stimulus. As shown in Fig. 3 (left), this resulted in a transient response, both on the column level and in the MEG. The evolution of the state variables showed positive deflections peaking in the 100-ms latency range and negative deflections at about 200 ms. The accompanying firing



FIG. 3. Response profiles of the 208 columns of the model to an isolated tone. The left and right figure columns show the profiles in the case of slowly-decaying and fast-decaying adaptation, respectively. (A) In the case of slowly-decaying adaptation, the evolutions of the 208 state variables following the presentation of the tone show both positive and negative deflections. (B) The firing rates of the columns peaked in the 24-210 ms range with a median of 90 ms, corresponding with the peak latency of the N1m response. (C) The adaptation terms reached their respective minima within 50 ms after stimulus onset and decayed towards their resting values over several seconds. (D) The magnetoencephalographic (MEG) response to the 50ms tone peaked with an amplitude of 120 at 90 ms, and thus resembles the auditory N1m response. The amplitude of the MEG is calculated as the sum of the excitatory inputs to the pyramidal cells, weighted by the synaptic weights and the suppression terms [Eqn. (3)]. (E-H) With fast-decaying adaptation, the state variables, the firing rates and the MEG response all resembled the case with slowly-decaying adaptation. However, the adaptation terms decayed to their resting state by about 600 ms.

rates increased transiently and peaked in the 50-150 ms range. More specifically, assuming a 10-ms signal delay from cochlea to cortex (Liegeois Chauvel et al., 1991), the earliest onset latencies in the core, belt and parabelt were 17, 33 and 51 ms, respectively. These agree well with non-invasive results from the human auditory cortex, where corresponding serial activation occurs in the 17-48 ms range (Inui et al., 2006). The mean delay between response onset and maximum firing rate was 32 ms. The MEG signal of the model comprised a transient response peaking at 90 ms. The depression terms decreased from the rest value of a = 1 to a = 0.4 in about 200 ms, and thereafter recovered towards a = 1 over several seconds. When the adaptation time constant was decreased from its default value of  $\tau_a = 1.6$  s to 100 ms (Fig. 3, right), similar response patterns to that of the default condition were observed, with the exception of the depression term that recovered to its resting value (a = 1) by about 600 ms.

## SSA to repeated stimuli

In general, stimulus repetition led to adaptation: diminished responses were observed, both on the column level and in the MEG, with the effect depending on SOA (Fig. 4A). As the SOA was decreased from 12 s to 200 ms, the peak firing rates averaged over the columns decreased from 0.18 to 0.006 (97% drop) in the case of slowly-decaying adaptation ( $\tau_a = 1.6$  s), and a similar decrease was observed for an intermediate value of the adaptation time constant ( $\tau_a = 400$  ms). A much smaller relative decrease was observed with fast-decaying adaptation ( $\tau_a = 100$  ms): from 0.2 to 0.16 (25% drop). Similarly, the peak amplitude of the N1m depended monotonically on the SOA. This dependence could be described by an exponentially saturating function similar to that used by Lu et al. (1992). The time constant of this gently increasing curve was 3.0 s for slowly-decaying adaptation. With  $\tau_a = 400 \text{ ms}$ and  $\tau_a = 100$  ms, the N1m increase was more abrupt, with time constants of 600 ms and 90 ms, respectively.

It was determined whether these response decrements associated with stimulus repetition reflected SSA. Adaptation to a repeated stimulus is stimulus-specific if it does not generalize to other stimuli, that is, if response recovery occurs to stimuli deviating from the repeated stimulus. The oddball paradigm was used to test for SSA (Ulanovsky *et al.*, 2003), specifically, by presenting the model with a series of 753-Hz 'standard' tones interspersed with 1054-Hz 'deviants' (P = 0.1). As shown in Fig. 4B, the proportion of columns displaying SSA depended on the adaptation time constant, growing monotonically from zero to 11% (N = 22) as  $\tau_a$  was increased from 100 ms to 1.6 s. This increase of SSA columns were largest in the core and belt: the proportions of SSA columns were largest in the core and belt, where they grew to 10% (N = 5) and 13% (N = 17), respectively, for  $\tau_a = 1.6$  s. No SSA columns were found in the parabelt.

As illustrated in Fig. 4C, the increase in the synaptic adaptation time constant was also reflected in adaptation of the MEG responses. With  $\tau_a = 100$  ms, the standard and deviant elicited nearidentical N1 responses. With  $\tau_a = 1.6$  s, the deviant elicited a much larger response (peak amplitude 76) than the standard (peak amplitude 30), the difference between the two peaking at 90 ms and 140 ms. This enhancement of the response associated with the deviant was due to two factors. First, the average peak firing rate  $f_{\text{max}}$  of the columns activated during the MEG deflection was 0.18 following the presentation of the standard and 0.37 following the deviant. Second, the number of activated columns ( $f_{max} > 0$ ) was higher for the deviant (N = 81, of which 63 selective to deviant) than for the standard (N = 42, of which 24 selective to standard). Columns exhibiting SSA were a subset of the columns responding to the deviant (N = 29, 36% of columns activated by deviant), suggesting that SSA on the column level does not alone account for amplitude differences (i.e. mismatch responses) found with oddball stimulation. Indeed, the majority of the columns producing the response to the deviant (56%; N = 45) were columns that were *selectively* activated by the deviant, remaining unresponsive to the standard  $(f_{\text{max}} = 0)$ both in the oddball and the isolated-standard condition (and thus not of the SSA kind).

# Responses to tone pairs at the column level

The model was presented with a stimulus set comprising four sounds: an ascending pair of tones with respective frequencies of 1054 Hz and 1476 Hz (tone duration = 50 ms, SOA 600 ms); the reversed, descending version of this tone pair; and the isolated



FIG. 4. Adaptation to repeated stimulation in the model. (A) On the left, the peak firing rates  $f_{max}$  averaged over the columns is a monotonically increasing function of stimulus-onset asynchrony (SOA). On the right, the peak amplitude of the N1m grows at a saturating rate as SOA is increased. Denoting the peak amplitude by *A*, this monotonic behaviour could be described by the function  $A(SOA) \propto [1 - exp(-SOA/\tau)]$ , where  $\tau$  is the time constant of saturation. This has the values 90 ms, 600 ms and 3.0 s for synaptic adaptation with time constants 100 ms, 400 ms and 1.6 s, respectively. (B) An example of stimulus-specific adaptation (SSA) is shown on the left: in this column, the 753-Hz tone presented in isolation elicited a prominent response peaking at  $f_{max} = 0.78$ . When this tone was presented as the standard stimulus in the oddball paradigm, a diminished response of  $f_{max} = 0.35$  was elicited. When a 1054-Hz deviant was delivered, the response recovered in strength and peaked with an amplitude of  $f_{max} = 0.72$ . Thus, the adaptation due to the repetition of the standard is stimulus-specific. On the right, the proportion of SSA columns were found in parabelt (bottom). (C) The effect of synaptic adaptation is visible in the magnetoencephalographic (MEG) responses elicited in the oddball paradigm. On the left, fast-decaying adaptation resulted in negligible differences between the averaged response to the standard and that elicited by the deviant. With slowly-decaying adaptation, the standard stimulus elicited a diminished response. The presentation of the deviant lead to response recovery in the 50–200 ms range, and the difference between the two responses peaked at 90 ms and 140 ms. On the right, the columns activated by the stimuli ( $f_{max} > 0.1$ ) in the case of slowly-decaying adaptation are shown on an area-column map. The standard activated far fewer columns than the deviant. A minority of activated columns displayed SSA, with the majority producing stimulus-specific responses to either

1054-Hz and 1476-Hz tones. The two tone pairs elicited transient responses ( $f_{\text{max}} > 0.1$ ) from 130 out of the 208 columns (63%). In general, when looking at how the first tone affected the response to

the second tone, both suppression and enhancement effects were observed (70% and 30% of cases, respectively). To determine whether these responses represent temporal integration, we examined

their CS, that is, how selective these responses were with regard to the entire stimulus set. We found that  $P_{rev} = 20\%$  (N = 41) columns preferred one or the other tone pair compared with the reversed version. Further,  $P_{cnt} = 13\%$  (N = 26) columns responded more strongly to a tone pair than to the constituent tones. The total number of CS columns, which responded only to a specific tone pair while remaining unresponsive to the other stimuli was  $P_{\rm CS} = 12\%$ (N = 24). This represented 10% of columns in the core, 12% of belt columns and 9% of parabelt columns. Examples of column activation displaying CS and of temporally non-selective columns are shown in Fig. 5. The dependence of this temporal integration on synaptic depression was tested by decreasing the depression time constant to  $\tau_a = 100$  ms. In this case, no CS columns were found, with the corresponding values being  $P_{rev} = 0$ ,  $P_{cnt} = 0$ ,  $P_{int} = 0$ . Thus, synaptic depression with the slower decay constant was necessary for temporal integration to occur.

Additional simulations confirmed that the emergence of CS did not depend on the exact parameter values of the model. Variations of the weight values in a 25–200% range showed that 23% of combinations resulted in a proportion of CS columns of 10% or higher. As shown in Fig. 6, this proportion depended non-monotonically on both the probability of intra-columnar connections ( $P_0$ ) and the probability of inhibitory connections ( $P_{inh}$ ). Values of  $P_{CS} > 10$ were found when  $P_0$  was in the 0.3–0.9 range, and  $P_{inh}$  was in the 0.4–0.7 range.

As shown in Fig. 7, the role of synaptic depression in the temporal integration of tone pairs with variable SOAs in the 50–2360 ms range was investigated. With slow decay of depression ( $\tau_a = 1.6$  s), the number of temporally sensitive columns  $P_{rev}$  decreased from 46% to 6% (N: 95–12) over the SOA range. With fast decay  $(\tau_a = 100 \text{ ms})$ , the number of reversal-sensitive columns sank from  $P_{\rm rev} = 46\%$  (N = 95) to zero by the inter-tone interval of 400 ms. Further, the proportion of columns selective to tone pairs vs. isolated tones peaked at  $P_{cnt} = 12\%$  (N = 24) at SOA = 575 ms for slow decay of depression, and thereafter slowly decreased as SOA was increased; in the case of fast decay,  $P_{cnt}$  was 9% (N = 19) at the 260-ms SOA and negligible at other SOAs. In the slow-decay case, the total number of columns displaying temporal integration had a non-monotonic relationship with SOA, increasing from 1% (N = 3) to a maximum of  $P_{\rm CS} = 11\%$  (N = 22) at the 575-ms SOA and thereafter slowly decreasing. With fast decay of adaptation, temporal integration occurred only at the 260-ms SOA at which  $P_{\rm CS} = 9\%$  (N = 18). With an intermediate value of  $\tau_{\rm a} = 400$  ms,  $P_{\rm CS}$  peaked at 260 ms with a value of 9% and decreased to a negligible level by 1030 ms. In all cases, there was a threshold of 260 ms below which only 0.5-1% (N: 1-3) columns exhibited temporal integration. In sum, temporal integration was weak below 260 ms, and slow decay of adaptation was necessary for temporal integration to occur at time spans beyond 400 ms.

Using the above tone-pair stimuli and varying the adaptation decay constant in the 0.1–1.6 s range, we looked at the distribution of CS columns in the different regions of the model. In general, the proportion of CS columns was the highest in the core and the lowest in the parabelt, although in some cases there was little difference between the three regions. Figure 8 shows the proportions of CS columns for the SOAs of 575 ms and 1030 ms. In both cases, the total proportion of CS columns in the model grew monotonically to about 10% as  $\tau_a$  was increased to 1.6 s. With the 575-ms SOA, the largest regional differences occurred with  $\tau_a = 0.8$  s, when  $P_{CS}$  for the core, belt and parabelt measured 17%, 8% and 3%, respectively.



FIG. 5. Demonstration of combination sensitivity (CS) in individual columns of the model. The figure shows examples from the activations (i.e. time course of firing rates) of 18 core and belt columns: the nine on the left show CS, whereas the nine on the right are non-selective to temporal order. (A) On the left, the nine columns responded to the second tone of a tone pair (onset at 600 ms; black rectangle in the stimulus diagram). (B–D) The tone pair in the reversed direction and the tones presented in isolation elicited no responses in these same columns. Thus, the columns respond selectively to the temporal structure of stimulation. (E–H) These nine columns generated responses in all conditions, and therefore are unselective to temporal structure. All examples are from the model with slowly-decaying adaptation.



FIG. 6. The dependence of combination sensitivity (CS) on the connection probabilities. The probability of intra-column connections ( $P_0$ ) and the probability of inhibitory connections ( $P_{inh}$ ) were each varied in the 0–1 range. The proportion of CS columns showed tuning to connection probability, having a non-monotonic dependence on both  $P_0$  and  $P_{inh}$ . The default values used in the model are represented by the star. The values represent means from 10 separate runs.

With the 1030-ms SOA,  $P_{CS}$  for the core was consistently larger than that for the parabelt, each growing monotonically as a function of  $\tau_a$ , and reaching the respective values of 13% and 6% at  $\tau_a = 1.6$  s.

The above conditions for temporal integration include those for response facilitation in the two-tone paradigm used in primate studies (Brosch et al., 1999), in which facilitation is observed when the response to the second tone in a pair is larger than that evoked by the tone when presented in isolation. For comparison with these previous results, the responses to the above ascending and descending tones pairs were analysed using measures employed in electrophysiological experiments. In the case of slowly-decaying adaptation, we found that 16% (N = 34) of columns exhibited response facilitation. As shown in Fig. 9, the magnitude of facilitation ranged from 110% to 640% (median 141%), the distribution peaking in the 120-140% range and then monotonically decreasing for larger SOAs (Fig. 9A). Forward facilitation occurred across the whole range of SOAs (50-2360 ms), with the distribution of SOAs evoking the strongest effect peaking at 50 ms and 785 ms (Fig. 9B). The longest SOA at which forward facilitation was observed also covered the whole range of SOAs, peaking at 85 ms and 2360 ms, with a median of 418 ms. In the case of fast-decaying adaptation, 16% (N = 33) columns showed facilitation ranging from 110% to 700% (median 143%), with a peak at 110% facilitation followed by a monotonic decrease. Facilitation was observed in the 50-575 ms range, and was a non-monotonic function of SOA: the SOA with the strongest facilitating effect as well as the longest SOA that produced facilitation both peaked at 260 ms.

# Simulated MEG responses to tone pairs

To demonstrate how synaptic adaptation is reflected non-invasively, the above 1054-1476 Hz and 1476-1054 Hz tone pairs (tone duration = 50 ms, SOA 600 ms) were presented in the oddball paradigm. The ascending tone pair was used as the standard (P = 0.9)



FIG. 7. Temporal integration of tone pairs at different stimulus-onset asynchronies (SOAs). (A) With the shortest-SOA tone pairs used as stimuli, about  $P_{\rm rev} = 45\%$  columns produced responses that were much stronger to the tone pair than to the reversed version of the tone pair.  $P_{\rm rev}$  decreased monotonically as SOA was increased. With fast-decaying adaptation ( $\tau_a = 100$  ms),  $P_{\rm rev}$  reached a negligible level at 400 ms SOA. With slowly-decaying adaptation ( $\tau_a = 1.6$  s),  $P_{rev}$  decreased to 6% at the longest SOA (2360 ms). (B) The proportion of columns  $(P_{cnt})$  producing responses to the tone pair but none to the individual tones presented in isolation had a non-monotonic relationship with SOA. In the fast-decaying-adaptation condition, a sharp peak was observed at SOA = 260. In contrast, with slowly-decaying adaptation, this measure had a broad distribution peaking at SOA = 575 ms. (C) The proportion of columns (Pint) satisfying both of the above conditions and thus demonstrating temporal integration also had a sharp and broad distribution for fast- and slowly-decaying adaptation, respectively. With an intermediate value of the adaptation constant ( $\tau_a = 400 \text{ ms}$ ), the various proportions behaved much like those measured in the fast-decay case.

and the descending tone pair as the deviant (P = 0.1). In both cases, the individual tones of the tone pairs elicited N1m-like responses, which all peaked with an amplitude of about 50, as shown in Fig. 10A. Thus, compared with the N1m elicited by isolated tones, peaking with an amplitude of about 120, both the standard and the deviant elicited diminished responses. The response elicited by the first tone of the deviant was diminished compared with that elicited by the first tone of the standard. In the 100-ms time range following the peak of the response to the second tone, the response to the deviant was enhanced compared with the response to the standard. The difference curve (derived by subtracting the standard response from that to the deviant) reached its maximum at 124 ms following the onset of the tone. As in the case of oddball stimulation using single tones (see Fig. 4C), this enhancement could be traced to two reasons: firstly, the average peak firing rate of activated columns was higher for the deviants ( $f_{max} = 0.33$ ) than for the standards  $(f_{\text{max}} = 0.29)$  and, secondly, the number of activated columns was higher for the deviant (P = 29%, of which 67% selective to deviant) than for the standard (P = 24%, of which 80% selective to standard). Crucially, when the adaptation time constant was decreased to 100 ms, the difference between the standard and deviant responses was abolished, as were the differences in average firing rate (0.49 and 0.50 for the standard and deviant, respectively) and activated columns (41% and 39%).



FIG. 8. Temporal integration of tone pairs as a function of the adaptation time constant. (A) When the tone pair had an stimulus-onset asynchrony (SOA) of 575 ms, the proportion of CS columns  $P_{\rm CS}$  grew monotonically as  $\tau_{\rm a}$  was increased. (B) With the adaptation constants below 1.6 s, the proportion of CS columns was larger in the core than in the belt. (C) The monotonic relationship between  $P_{\rm CS}$  and  $\tau_{\rm a}$  was present also with the longer SOA of 1030 ms. (D) The proportion of CS cells was consistently highest in the core and lowest in the parabelt.

Figure 10 demonstrates further examples of N1m variations (i.e. mismatch responses) elicited by tone pairs. Tone pairs with an SOA of 50 ms were presented in the oddball paradigm. This time, the stimuli elicited unimodal deflections, with the deviant resulting in a larger response than the standard (Fig. 10B). Decreasing  $\tau_a$  to 100 ms abolished this difference in the responses. We also employed a stimulation paradigm used to demonstrate MMN to 'abstract' stimulus features (Korzyukov et al., 2003). In this case, a set of four different ascending tone pairs was presented equiprobably as the standards, and the deviants were drawn equiprobably from a set of four corresponding descending tone pairs. The purported abstract feature that separates the standards from the deviants is the direction of the frequency step. Again, as shown in Fig. 10C, the result was a prominent difference between the standard and deviant responses and, again, the employment of fast-decaying synaptic depression abolished this difference. In fact, the model predicts that any arbitrary division of multiple tone pairs into standard and deviant categories results in an amplitude differential between the responses to the standards and deviants. An example of this is presented in Fig. 10D, where the standards comprised both ascending and descending tone pairs at different frequencies, and the deviants were the time-reversed versions of the standards. Once more, the deviants produced a response enhancement. In sum, slowly decaying adaptation seems to be necessary for differential responses to occur in the oddball paradigm.

In view of the debate on the origin of the mismatch response (Näätänen *et al.*, 2005), we ascertained whether some of the columns responding to the deviant constituted a separate MMN generator, that is, a set of columns that produce the difference curve but that remain unresponsive to isolated stimuli (Näätänen, 1992; May & Tiitinen, 2010). For this, we mapped those columns (designated



FIG. 9. Population characteristics of forward facilitation in the model. The results for slowly- and fast-decaying adaptation ( $\tau_a = 1.6$  s and 100 ms, respectively) are shown on the left and right, respectively. (A) The magnitude of forward facilitation was the ratio between the maximum response  $f_{\text{max}}$  to the second tone of a tone pair and  $f_{\text{max}}$  to this same tone when presented in isolation. The distribution peaked in the 120–140% range for slowly-decaying adaptation. (B) The distribution of the stimulus-onset asynchrony (SOA) that produced maximal facilitation peaked at 50 ms. (C) The longest facilitation gale at 85 ms and 2360 ms. (D) For fast-decaying adaptation, the magnitude of facilitation peaked in the 110–120% range. (E) The SOA evoking strongest facilitation peaked at 260 ms. (F) The longest SOA producing facilitation peaked at 260 ms.

*C*) that were activated ( $f_{\text{max}} > 0$ ) by a white noise burst, by a complex tone, or by any of 16 isolated tones in the 100–15 557 Hz range. In each of the oddball experiments, all the columns responding to the deviants ( $f_{\text{max}} > 0$ ) belonged to the set *C*. This means that all of the columns contributing to the deviant response, and therefore to the MMN difference curve, also responded to isolated stimuli. Thus, while the model produced prominent mismatch responses, there was no separate 'mismatch response generator' activated exclusively by stimulus change.

# Single-column and MEG responses to complex sounds

In these two experiments, we investigated how the temporal integration of complex tone sequences is supported by synaptic adaptation.



In the first experiment (see Fig. 2B), the presence of CS responses was investigated by presenting 70 distinct sequences of four tones. In each case, the response to the sequence was contrasted to the respective responses elicited by the time-reversed version of the sequence and the separate stimulus halves, that is, rump sequences comprising either the first and second tones, or the third and fourth tones. As in the tone-pair experiments, columns exhibiting CS were identified as those producing stronger responses (> 100% increase)

FIG. 10. Magnetoencephalographic (MEG) responses of the model in the oddball paradigm. (A) An ascending tone pair with an inter-tone stimulusonset asynchrony (SOA) of 600 ms was used as the standard stimulus, and the corresponding descending tone pair was presented as the deviant. With slowly-decaying adaptation ( $\tau_a = 1.6$  s, left), the response to the second tone of the deviant pair was enhanced in the 100-200 ms post-stimulus range compared to the equivalent response to the standard. Fast-decaying adaptation ( $\tau_a = 100$  ms, right) abolished this difference. (B) When the SOA of the tone pairs was reduced to 50 ms, an even more marked difference developed between the responses to the standard and deviant. The difference wave peaked at the NIm peak latency and again at 150 ms. Reducing the adaptation time constant resulted in near-identical responses to the standard and deviant. (C) The model was presented with ascending and descending tone pairs as the standard and deviant, respectively. The tone pairs, occurring at multiple frequencies, were distinguishable from each other only by the 'abstract' feature of the direction of the frequency step. Slowly-decaying adaptation resulted in a prominent enhancement of the deviant response in the 80-150 ms range, and fast-decaying adaptation lead once more to responses to the deviant that were near-identical to the standard responses. In sum, these simulations demonstrate that a differential response in the oddball paradigm to tone pairs requires adaptation with a long time constant. (D) The model predicts that any combination of tone pairs into standard and deviant categories (i.e. any arbitrary 'abstract' rule, including the one above) results in mismatch responses. This example utilized the same experimental set-up as above, except that the standard category comprised both ascending and descending tone pairs, as did the deviant category, where each deviant tone pair was a reversed version of a standard pair. The schematic diagram on the left shows the collection of tone pairs used as the standard and those used as the deviant. A prominent mismatch response resulted from this arbitrary division of stimuli ( $\tau_a = 1.6$  s). Note that in the cases of slowly-decaying adaptation, the responses to the deviants are all diminished with respect to the peak amplitude of the response elicited by an isolated tone ( $R \sim 120$ ; see Fig. 3D).

to the forward sequence than to any of the other three stimuli. As depicted in Fig. 11A, the number of the CS columns, averaged over the 70 sequences, was found to be a monotonically increasing function of the adaptation time constant. When  $\tau_a$  was increased from 0.1 to 1.6,  $P_{\rm CS}$  rose from a value of 2% to 11%. Thus, lengthening the time course of synaptic depression was directly reflected in the ability of the model to perform temporal integration of complex tone sequences. This monotonic relationship between the adaptation time constant and  $P_{\rm CS}$  was seen in each of the regions of the model (Fig. 11B). Interestingly,  $P_{\rm CS}$  was consistently higher in the belt and parabelt (where it increased to about 12%) than in the core (where it reached 6%).

In the second experiment (see Fig. 2C), the model was presented with the 1054-Hz tone in isolation as well as with 24 five-tone sequences that all comprised the same tones but, in each sequence, the order of these tones was unique, and the final tone was always of 1054-Hz frequency. In correspondence with forward facilitation experiments using tone pairs, we examined whether the tone combinations enhanced the response to the final tone. Of particular interest were cases where only one of the combinations resulted in substantial forward facilitation (i.e. a 100% increase in response strength). As shown in Fig. 12A, the proportion  $P_{\text{fac}}$  of such columns, highly selective to temporal structure, was found to depend on the adaptation time constant. With the shortest value of  $\tau_a$  (0.1 s), the number of selective columns was 11% (N = 23). As  $\tau_a$  was increased to 0.8 s, the proportion  $P_{\text{fac}}$  grew to 22% (N = 45), and remained at this level for further increases of  $\tau_a$ . Thus, columns exhibited responses that were facilitated by specific temporal combinations of sounds, and the number of these columns had a monotonic relationship with the adaptation time constant. The region-specific  $P_{\text{fac}}$  values resembled that of  $P_{\text{fac}}$  for the total proportion of columns, each one plateauing out at  $\tau_a = 400$  ms. The proportion of columns exhibiting facilitation was consistently highest in the parabelt. The





Adaptation time constant  $\tau_a$  (s)

FIG. 11. Combination sensitivity (CS) of responses to complex tone sequences. (A) When sequences of four different tones were used as stimuli, the proportion  $P_{\rm CS}$  of the total number of columns exhibiting CS increased monotonically as a function of the adaptation time constant  $\tau_{\rm a}$ . CS was achieved when the response to a sequence was at least double in magnitude than: (a) the response to the time-reversed sequence; (b) the response to the initial half of the sequence presented in isolation; and (c) the response to the isolated second half. (B) The proportion of CS columns was consistently lowest in the core. The figure shows the mean values of  $P_{\rm CS}$  from the set of 70 unique sequences. Bars denote standard error of the mean (SEM).

maximum values of  $P_{\rm fac}$  for the core, belt and parabelt were 20%, 22% and 25%, respectively.

Selectivity to the temporal structure of complex tone sequences could be made visible in the MEG by using the oddball paradigm (Fig. 13). Two sequences from the second experiment, each comprising an identical set of five tones, were used. The second sequence differed from the first only in that the third and fourth tones were in reverse order. When the first and second sequences were presented as the standard and deviant stimulus, respectively, there was a marked increase in the magnitude of the response to the deviant compared with that elicited by the standard. This enhancement (from 22 to 44, 114%) occurred in a 150-ms time window following the presentation of the third tone. It was associated with stronger firing rates at the column level (mean 0.2 vs. 0.3 for standard and deviant, respectively), an increase in activated columns (12% vs. 17%) and a shift in the spatial distribution of activity. The differential response was abolished by decreasing  $\tau_a$  to 100 ms.

# Discussion

The temporal integration of auditory information is reflected by the CS of auditory cortex responses measured in single cells (McKenna *et al.*, 1989; Lewicki & Konishi, 1995; Wang *et al.*, 1995; Brosch

FIG. 12. Selective facilitation of responses to complex tone sequences. (A) In this experiment, the facilitating effect of tone sequences on the response to a subsequent tone was evaluated. A set of 24 five-tone sequences was presented to the model. Each sequence was a unique combination of the same five tones, and the final tone was always of frequency 1054 Hz. The responses to this final tone were contrasted to each other and to the tone presented in isolation. Facilitation indicating temporal integration was judged to occur when a column responded selectively to only one of the sequences (i.e. producing a response at least double in magnitude than the responses to the other sequences and the response to the isolated tone). This kind of selective facilitation was enhanced by increasing the time constant of adaptation. (B) The proportion of columns exhibiting facilitation was always the largest in the parabelt. The values in the figure represent the number of columns  $P_{\rm fac}$  exhibiting facilitation averaged over 70 runs (see Materials and methods). Bars denote SEM.

& Schreiner, 1997, 2000; Rauschecker, 1997; Brosch et al., 1999; Brosch & Scheich, 2008; Recanzone, 2008; Sadagopan & Wang, 2009). It is also suggested by the non-invasively measured N1m, which varies in amplitude distribution according to the historical context of the eliciting stimulus. Here, it was demonstrated that these two phenomena can be explained by adaptation mediated by activity-dependent synaptic depression, which in this study was restricted to the excitatory connections between the pyramidal cell populations. A computational model of auditory cortex that included a serial core-belt-parabelt structure of multiple parallel streams of activation was used (May & Tiitinen, 2013). By varying the time course of adaptation, we found responses indexing the temporal integration of tone pairs and longer, complex sequences of tones. Specifically, responses exhibited both facilitation by and selectivity to the temporal structure of stimulation. The model also produced various adaptation phenomena, namely SSA responses in single-column measurements and variations in N1m responses (i.e. 'MMN' responses) obtained in the oddball paradigm. Importantly, the prevalence of responses reflecting temporal integration and SSA increased



FIG. 13. Magnetoencephalographic (MEG) responses to complex tone sequences presented in the oddball paradigm. (A) Two sequences of five tones used in the forward facilitation experiment were used as respective standard and deviant stimuli. As shown in the stimulus diagram (bottom), the deviant comprised the same set of tones as the standard, but with the third and fourth tone in reverse order. With slowly-decaying adaptation, this resulted in prominent differences between the standard and deviant responses starting after the presentation of the third tone (top). Fast-decaying adaptation abolished these relative differences in response amplitude. (B) The maximum firing rates  $f_{\text{max}}$  of the columns during 80–150 ms following the onset of the third tone are shown for the case of slowly-decaying adaptation. Compared with the activity evoked by the standard, the columns respond more vigorously to the deviant, activity is more widespread and there is a difference in the spatial distribution of activation. This shift in the 'centre of gravity' of activity would lead to a shift in the estimated source location if a point-like ('single-dipole') source model was used in the estimation.

monotonically as a function of the time constant of synaptic adaptation. To replicate the adaptation of the N1m response, the long time constant of 1.6 s was required.

A general feature that emerged in the simulations was that the complexity of the stimulation seemed to be reflected in the regional distribution of columns exhibiting temporal integration. With the two-stimulus oddball paradigm, SSA columns were found exclusively in the core and belt regions, with none in the parabelt (Fig. 4). Similarly, the highest and lowest proportions of columns showing temporal integration of tone pairs tended to be in the core and parabelt, respectively (Fig. 8). The situation was reversed when four- and five-tone sequences were used: the proportion of columns producing responses indicating temporal integration was always lowest in the core and higher in the belt and parabelt (Figs 11 and 12). Thus, while needing further confirmation, the current results suggest that serial structure boosts the ability of auditory cortex to represent temporally complex sounds: the more temporal complexity a sound

has, the deeper up the serial structure its representation as a temporal whole seems to be moved.

# The temporal integration of tone pairs

The ability of the model to perform temporal integration of tone pairs was examined. Integration was judged to have occurred if columns of the model responded stronger to a tone pair in the forward direction than in the reversed direction. A further condition was that neither of the tones presented in isolation elicited a robust response. A column satisfying these conditions was CS, and was therefore guaranteed to respond strongly to the second tone if, and only if, it was preceded by the first tone, and therefore its response indicated that temporal integration of the tone pair had occurred. We found that adaptation with a fast decay constant of 100 ms supported temporal integration of 260-ms tone pairs in 9% of columns. For pairs extending beyond this limit, the slower decay constant of 1.6 s was required. In this case, a maximum of 12% of columns exhibited CS when the SOA of the tone pair was 575 ms. Importantly, slower decay of adaptation allowed for temporal integration to occur at all the SOAs tested (up to 2360 ms). Interestingly, there was a threshold of integration in the sense that there was little evidence for it at the shortest SOAs (< 260 ms) using the stringent requirements of the current analyses (i.e. > 100% facilitating effects on response amplitude).

The current results resemble those of Lewicki & Arthur (1996), who looked at the temporal CS of responses in the songbird forebrain. Using song syllables as stimulus material, they found that in area HVc, 27% of units responded stronger to a two-syllable sequence than to its reversed version (equivalent to our  $P_{rev}$ ); 19% showed stronger responses to the sequence than to the individual syllables presented in isolation (equivalent to  $P_{cnt}$ ); and 19% exhibited a combination of these preferences (equivalent to  $P_{CS}$ , i.e. our CS requirements). Our simulations (Fig. 7) mirror this pattern in two respects. First, for all tone-pair durations and adaptation time constants,  $P_{\rm rev}$  was always larger than  $P_{\rm cnt}$ . Second,  $P_{\rm CS}$  was very close to P<sub>cnt</sub> and in many cases identical to it. For example, with the 400-ms adaptation time constant and the 260-ms SOA (i.e. 310 ms sequence duration, corresponding with the approximate 300ms duration of the syllable sequence of Lewicki & Arthur, 1996), we measured  $P_{\rm rev} = 33\%$ ,  $P_{\rm cnt} = 9\%$  and  $P_{\rm CS} = 9\%$ . Furthermore, the presence of CS responses correlates with the results from monkey. Rauschecker (1997) observed that complete monkey vocalizations produce larger responses than individual segments of the call presented in isolation. Preference to forward vs. reversed speciesspecific vocalizations has been observed in monkey auditory cortex (Wang & Kadia, 2001; Sadagopan & Wang, 2009), with the reported proportion of such cells varying from 10% (Recanzone, 2008) to 75% (Wang et al., 1995). In the current simulations, preference to forward vs. reversed tone pairs was found in 46% of columns at and below inter-tone intervals (SOAs) of 155 ms. At longer intervals, preference declined as a function of interval and depended on the adaptation time constant. Thus, the current observations suggest that the number of CS cells and those showing sensitivity to sequence direction depends on stimulus duration and variations in the adaptation time course.

A second measure of temporal integration looked at forward facilitation, where the response to a tone is enhanced if it is preceded by another tone. This is in fact a subset of the conditions for CS as defined above, and it was used so that comparisons could be made to previous animal studies of two-tone facilitation in the primate auditory cortex (Brosch *et al.*, 1999; Brosch & Schreiner, 2000;

# 628 P. J. C. May et al.

Bartlett & Wang, 2005; Brosch & Scheich, 2008; Sadagopan & Wang, 2009). With both fast- and slowly-decaying synaptic adaptation, 16% of columns in the model showed facilitation - a lower proportion than the 40-60% values found in single-unit measurements in primates (Brosch & Schreiner, 2000; Bartlett & Wang, 2005; Brosch & Scheich, 2008). Irrespective of the time constant of synaptic adaptation, the magnitude of facilitation peaked in the lower range (110-140%), resembling the results from primate studies. This could also be said of the distribution of the preferred SOA, which was weighted at the lowest SOAs. However, there was a clear difference between the two time constant conditions in that facilitation with fast-decaying adaptation was observable only at SOAs below 575 ms (median 260 ms), whereas facilitation was present over the whole SOA range with slowly-decaying adaptation. In this sense, the results with fast-decaying adaptation are more closely aligned with results from primates (e.g. longest facilitating SOAs 539 ms, median 187 ms; Brosch & Scheich, 2008). This leads to tentative predictions, based on the link between N1m amplitude and adaptation decay. First, our results suggest that the N1m equivalent in primates could have a faster saturation as a function of SOA than does the N1m in humans. Second, invasive measurements of facilitation should produce effects that persist at longer SOAs in humans than in non-human primates.

### The temporal integration of complex tone sequences

The temporal integration of complex tone sequences was studied through implementing the CS and forward facilitation measures used in the previous tone-pair paradigms. In the case of four-tone sequences, we found a monotonic relationship between the adaptation time constant and the number of columns exhibiting CS, which increased from 2% to 11% as  $\tau_a$  grew from 100 ms to 1.6 s. Using a modified version of forward facilitation measurements, sharp selectivity to the temporal structure of tone combinations was discovered. The model was presented with sets of tone sequences where each one was a unique combination of the same set of tones. Thus, the same spectral information was delivered in multiple temporal combinations. Several columns turned out to be selective to only one of these combinations, thus showing that temporal integration was taking place. The proportion of these columns (measured by  $P_{\rm fac}$ ) doubled from 11% to 22% as the adaptation decay constant was increased from 100 ms to 0.4 s and remained at this level for further increases of the constant. This saturation of  $P_{\rm fac}$  occurred at about the 650-ms duration of the tone sequences, indicating that the adaptation constant is directly reflected in the time window over which temporal integration of complex sequences occurs. The results of these two experiments offer experimentally verifiable predictions, and suggest that responses pointing to temporal integration of complex sequences (using the current measures) are equally prevalent as two-tone facilitation (16%, see above). In sum, these observations point to synaptic adaptation enhancing the ability of auditory cortex to perform temporal integration.

The fact that so many columns in the five-tone stimulation condition displayed integration even at the shortest adaptation decay constant might be due to the short, 150-ms inter-tone SOA used in the experiment. The silent 100-ms interval between the tones would have been too short to allow the dynamic variables of the model – the state variables u(t) as well as the adaptation terms a(t) – to reach their resting states. Thus, at the onset of each new tone, the response was determined not only by the afferent stimulation but also by the dynamic, non-resting state of the model. This mode of state-dependent computation provides an effective mode for making non-linear mappings of temporal information, as described in the theory of echo-state networks (Lukoševicius & Jaeger, 2009; see also Buonomano & Maass, 2009). Thus, the current model can also be regarded as a non-linear mapping of the inputs into a 208-dimensional space, where temporal integration causes the same input to be represented with different trajectories depending on an 'echo' resulting from previous inputs. In this context, the current results point to the possibility that echo-state processing might benefit from adaptation type of dynamics.

#### The role of adaptation in MEG responses

The model reproduced a range of non-invasively measured phenomena of the activation of auditory cortex. Upon stimulus presentation, the model generated a prominent transient MEG deflection that resembled the N1m in that: (1) the deflection peaked at about 100 ms post-stimulus; and (2) the amplitude of this deflection was dependent on the stimulus repetition rate, increasing monotonically as a function of SOA. This dependence on the SOA could be modelled with an exponentially saturating function used to describe the N1m measured in humans. With the default  $\tau_a$  of 1.6 s, the time constant of this saturation was about 3 s, which is within the 0.5– 5 s range measured by Lu *et al.* (1992). The use of the shorter decay time of synaptic adaptation ( $\tau_a = 100$  ms) resulted in an unrealistic, much smaller variation in the N1m amplitude.

A variety of differential 'mismatch' responses was generated when temporally structured stimuli were presented in the oddball paradigm. Prominent differences between the MEG responses to standards and deviants were observed when stimulation comprised tone pairs (both with long and short inter-tone SOAs), multiple tone pairs differing in the 'abstract' feature of tone step direction, and complex sequences of tones with differing temporal order. These amplitude variations resemble experimental results to tone pairs and sequences (for reviews, see Picton et al., 2000; Näätänen et al., 2010). However, the current model predicts that an amplitude enhancement of the response to the deviant results from any arbitrary rule for dividing multiple different tone pairs into standards and deviants. That is, the 'abstract' rule of tone step direction, used in experiments demonstrating 'primitive intelligence' (Näätänen et al., 2001; Korzyukov et al., 2003), just happens to be one of them. Thus, it is possible that the auditory cortex, rather than processing assumed abstract features of stimulation, is merely integrating stimulus events and responding to them selectively. In all the oddball experiments reported here, shortening the decay constant of synaptic depression abolished these mismatch responses. In sum, adaptation realised through persistent synaptic adaptation was necessary for the model to generate MEG responses consistent with experimental results from humans.

The variations in the amplitude distributions of the MEG responses obtained with oddball stimulation could consistently be explained by the following factors. First, the repeated presentation of the standard lead to response adaptation, a diminishing of response amplitude compared with responses elicited by the isolated standard. Second, the deviant resulted in response recovery, an enhancement of response amplitude. This was associated with an increase in the firing rates of the responding columns as well as with an increase in the number of columns responding to the deviant. We also noted shifts in the spatial distribution of activity associated with the presentation of the deviant. It has previously been suggested that SSA observed in single-cell experiments could be a correlate of the mismatch response (Ulanovsky *et al.*, 2003, 2004). In the current model, SSA provided part of the explanation. We found that only a

third of the columns contributing to the response to the deviant were responding in the SSA framework (i.e. producing responses to the standard that were diminished compared with the responses elicited by the deviant and with those elicited by the standard presented in isolation), with a majority of columns representing selective, rather than recovered, responses to the deviant. In the confines of the present modelling work, we found no evidence for mismatch-response generators, that is, columns that would respond only when a stimulus represents a change in repetitive stimulation (as suggested by Näätänen, 1992). Overall, variations in the N1m response elicited by oddball stimulation are explained by a combination of SSA and stimulus selectivity on the column level. Importantly, both of these factors require slowly-decaying synaptic adaptation.

# Temporal integration and adaptation: a link between intracortical and non-invasive responses

Adaptation in cortex is usually described in terms of forward masking effects whereby the response to a probe tone is suppressed if it is preceded by a masker tone (Calford & Semple, 1995; Brosch & Schreiner, 1997, 2000; Bartlett & Wang, 2005; Brosch & Scheich, 2008). A direct extension of this phenomenon is the suppression of the response to a repetitive stimulus, where each stimulus in effect masks the responses to subsequent stimuli (Ulanovsky et al., 2003). Previous MEG studies have suggested that adaptation would also explain N1m response recovery in the oddball paradigm, that is, when repetitive stimulation is interrupted by novel stimulation that then activates non-adapted neural populations (Butler, 1968; May et al., 1999; Ulanovsky et al., 2003; Jääskeläinen et al., 2004; May & Tiitinen, 2010). Indeed, the key role of adaptation in information processing in cortex is currently thought to be that of change detection and representation of stimulus statistics (for recent reviews, see Malmierca et al., 2014; Pérez-González & Malmierca, 2014).

In contrast to the above view and in concordance with previous computational studies (Buonomano & Merzenich, 1995; May & Tiitinen, 2007, 2010; Buonomano & Maass, 2009), the current results demonstrate that synaptic adaptation might have an unexpected function in contributing to temporal integration of auditory information - possibly the central function of auditory cortex (Nelken, 2004) - by allowing for individual cortical neurons to respond selectively to the temporal structure of stimuli. A counterintuitive consequence of this is that activity-dependent depression of excitatory synapses between pyramidal cell populations not only leads to forward suppression in the case of masker-tone interactions, but could also underlie response facilitation reflecting temporal selectivity (see also Goudar & Buonomano, 2014). This can be understood on the system level: by modifying the interaction weights between cortical cell populations in a stimulus-specific way, adaptation changes the input-output mappings between the areas of auditory cortex, making these depend on stimulation history. Thus, the response of an individual cortical unit to an incoming stimulus can either be suppressed or enhanced, depending on the immediate historical context of the stimulus. This explanation of response dependence on historical context bypasses the need to postulate specific circuitry, delay lines or learning: selectivity to the temporal structure of any stimulus arises stochastically in some subset of neurons, given a large enough network and synaptic plasticity (Lee & Buonomano, 2012).

The current results suggest that the mechanism underlying CS and forward facilitation in single-unit responses is the same mechanism that allows for change detection responses measured in singleunit activity (i.e. SSA) and non-invasively in the MEG and EEG. The current results point to the possibility that adaptation not only forms a memory system (in addition to the memory effects observed in cortical spike trains; Campbell *et al.*, 2010; Klampfl *et al.*, 2012), but also contributes directly to stimulus selectivity necessary for repetitive, temporally structured stimuli to be differentiated from novel events. Thus, our results suggest that there may be a link between temporal integration as reflected in *response selectivity* of cortical units and the context sensitivity of non-invasively measured MEG responses where temporal integration shows up as *response variations*, that is, the dependence of the N1m amplitude on the historical context of the eliciting stimulus.

It remains to be seen how synaptic adaptation might support the temporal integration used in more ecologically valid settings. The current simulations were limited to responses produced by isolated tone sequences and tones within stationary oddball distributions. In contrast, natural sound environments contain multiple, continuous, complex streams of sound, and the segregation of these depends on the allocation of attention (Bregman, 1990). Thus, the role of synaptic adaptation in integrating continuous auditory streams in a noisy environment remains an open question. Modelling these processes of ASA will probably require a more detailed description of the auditory system, including that of the precortical pathway and of the way auditory cortex is modulated by top–down effects originating from beyond the parabelt.

# Acknowledgements

This study was supported by the Academy of Finland (project 257811), the Emil Aaltonen Foundation, and Svenska kulturfonden, Finland.

#### Abbreviations

ASA, auditory scene analysis; CS, combination sensitivity; EEG, electroencephalography; IPI, inter-pair interval; MEG, magnetoencephalography; MMN, mismatch negativity; SOA, stimulus-onset asynchrony; SSA, stimulus-specific adaptation.

# References

- Abolafia, J.M., Vergara, R., Arnold, M.M., Reig, R. & Sanchez-Vives, M.V. (2011) Cortical auditory adaptation in the awake rat and the role of potassium currents. *Cereb. Cortex*, **21**, 977–990.
- Bartlett, E.L. & Wang, X. (2005) Long-lasting modulation by stimulus context in primate auditory cortex. J. Neurophysiol., 94, 83–104.
- Bleeck, S., Sayles, M., Ingham, N.J. & Winter, I.M. (2006) The time course of recovery from suppression and facilitation from single units in the mammalian cochlear nucleus. *Hearing Res.*, 212, 176–184.
- Bregman, A. (1990) Auditory Scene Analysis. MIT Press, New York.
- Brosch, M. & Scheich, H. (2008) Tone-sequence analysis in the auditory cortex of awake macaque monkeys. *Exp. Brain Res.*, 184, 349–361.
- Brosch, M. & Schreiner, C.E. (1997) Time course of forward masking tuning curves in cat primary auditory cortex. J. Neurophysiol., 77, 923–943.
- Brosch, M. & Schreiner, C.E. (2000) Sequence sensitivity of neurons in cat primary auditory cortex. *Cereb. Cortex*, 10, 1155–1167.
- Brosch, M., Schulz, A. & Scheich, H. (1999) Processing of sound sequences in macaque auditory cortex: response enhancement. J. Neurophysiol., 82, 1542–1559.
- Budd, T.W., Barry, R.J., Gordon, E., Rennie, C. & Michie, P.T. (1998) Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness. *Int. J. Psychophysiol.*, **31**, 51–68.
- Buonomano, D.V. & Maass, W. (2009) State-dependent computations: spatiotemporal processing in cortical networks. *Nat. Rev. Neurosci.*, 10, 113–125.
- Buonomano, D.V. & Merzenich, M.M. (1995) Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, 267, 1028–1030.
- Butler, R.A. (1968) Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. J. Acoust. Soc. Am., 44, 945– 950.

#### 630 P. J. C. May et al.

- Calford, M.B. & Semple, M.N. (1995) Monaural inhibition in cat auditory cortex. J. Neurophysiol., 73, 1876–1891.
- Campbell, R.A., Schulz, A.L., King, A.J. & Schnupp, J.W. (2010) Brief sounds evoke prolonged responses in anesthetized ferret auditory cortex. *J. Neurophysiol.*, **103**, 2783–2793.
- Douglas, R.J., Koch, C., Mahowald, M., Martin, K.A.C. & Suarez, H.H. (1995) Recurrent excitation in neocortical circuits. *Nature*, 269, 981–985.
- Garrido, M.I., Kilner, J.M., Kiebel, S.J. & Friston, K.J. (2007) Evoked brain responses are generated by feedback loops. *Proc. Natl. Acad. Sci. USA*, 104, 20961–20966.
- Goudar, V. & Buonomano, D.V. (2014) A model of order-selectivity based on dynamic changes in the balance of excitation and inhibition produced by short-term synaptic plasticity. *J. Neurophysiol.*, doi:10.1152/jn.00568.2014. [Epub ahead of print].
- Hackett, T.A., Stepniewska, I. & Kaas, J.H. (1998) Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J. Comp. Neurol., 394, 475–495.
- Inui, K., Okamoto, H., Miki, K., Gunji, A. & Kakigi, R. (2006) Serial and parallel processing in the human auditory cortex: a magnetoencephalographic study. *Cereb. Cortex*, 16, 18–30.
- Jääskeläinen, I.P., Ahveninen, J., Bonmassar, G., Dale, A.M., Ilmoniemi, R.J., Levänen, S., Lin, F.H., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H. & Belliveau, J.W. (2004) Human posterior auditory cortex gates novel sounds to consciousness. *Proc. Natl. Acad. Sci. USA*, **101**, 6809–6814.
- Kaas, J.H. & Hackett, T.A. (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. USA*, **97**, 11793– 11799.
- Klampfl, S., David, S.V., Yin, P., Shamma, S.A. & Maass, W. (2012) A quantitative analysis of information about past and present stimuli encoded by spikes of A1 neurons. *J. Neurophysiol.*, **108**, 1366–1380.
- Koch, C., Rapp, M. & Segev, I. (1996) A brief history of time (constants). *Cereb. Cortex*, 6, 93–101.
- Korzyukov, O.A., Winkler, I., Gumenyuk, V.I. & Alho, K. (2003) Processing abstract auditory features in the human auditory cortex. *NeuroImage*, 20, 2245–2258.
- Kurt, S., Deutscher, A., Crook, J.M., Ohl, F.W., Budinger, E., Moeller, C.K., Scheich, H. & Schulze, H. (2008) Auditory cortical contrast enhancing by global winner-take-all inhibitory interactions. *PLoS One*, 3, e1735.
- Lee, T.P. & Buonomano, D.V. (2012) Unsupervised formation of vocalization-sensitive neurons: a cortical model based on short-term and homeostatic plasticity. *Neural Comput.*, 24, 2579–2603.
- Levy, R.B. & Reyes, A.D. (2012) Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. J. Neurosci., 32, 5609–5619.
- Lewicki, M.S. & Arthur, B.J. (1996) Hierarchical organization of auditory temporal context sensitivity. J. Neurosci., 16, 6987–6998.
- Lewicki, M.S. & Konishi, M. (1995) Mechanisms underlying the sensitivity of songbird forebrain neurons to temporal order. *Proc. Natl. Acad. Sci.* USA, 92, 5582–5586.
- Liegeois Chauvel, C., Musolino, A. & Chauvel, P. (1991) Localization of the primary auditory area in man. *Brain*, 114, 139–151.
- Lu, Z.-L., Williamson, S.J. & Kaufman, L. (1992) Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Res.*, 572, 236–241.
- Lukoševicius, M. & Jaeger, H. (2009) Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.*, **3**, 127–149.
- Malmierca, M.S., Sanchez-Vives, M.V., Escera, C. & Bendixen, A. (2014) Neuronal adaptation, novelty detection and regularity encoding in audition. *Front. Syst. Neurosci.*, 8, 111.
- Markram, H., Wang, Y. & Tsodyks, M. (1998) Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. USA*, 95, 5323–5328.
- May, P.J.C. & Tiitinen, H. (2007) The role of adaptation-based memory in auditory cortex. *Int. Congr. Ser.*, 1300, 53–56.
- May, P.J.C. & Tiitinen, H. (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47, 66–122.
- May, P.J.C. & Tiitinen, H. (2013) Temporal binding of sound emerges out of anatomical structure and synaptic dynamics of auditory cortex. *Front. Comput. Neurosci.*, 7, 152.
- May, P., Tiitinen, H., Ilmoniemi, R.J., Nyman, G., Taylor, J.G. & Näätänen, R. (1999) Frequency change detection in human auditory cortex. J. Comput. Neurosci., 6, 99–120.

- McKenna, T.M., Weinberger, N.M. & Diamond, D.M. (1989) Responses of single auditory cortical neurons to tone sequences. *Brain Res.*, 481, 142– 153.
- Moeller, C.K., Kurt, S., Happel, M.F. & Schulze, H. (2010) Long-range effects of GABAergic inhibition in gerbil primary auditory cortex. *Eur. J. Neurosci.*, **31**, 49–59.
- de la Mothe, L.A., Blumell, S., Kajikawa, Y. & Hackett, T.A. (2006) Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *J. Comp. Neurol.*, **496**, 27–71.
- Näätänen, R. (1992) Attention and Brain Function. Erlbaum, Hillsdale, NJ.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. (2001) "Primitive intelligence" in the auditory cortex. *Trends Neurosci.*, 24, 283–288.
- Näätänen, R., Jacobsen, T. & Winkler, I. (2005) Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology*, **42**, 25–32.
- Näätänen, R., Astikainen, P., Ruusuvirta, T. & Huotilainen, M. (2010) Automatic auditory intelligence: an expression of the sensory-cognitive core of cognitive processes. *Brain Res. Rev.*, 64, 123–136.
- Nelken, I. (2004) Processing of complex stimuli and natural scenes in the auditory cortex. *Curr. Opin. Neurobiol.*, 14, 474–480.
- Nelken, I. (2014) Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biol. Cybern.*, **108**, 655– 663.
- Okada, Y.C., Wu, J. & Kyuhou, S. (1997) Genesis of MEG signals in a mammalian CNS structure. *Electroen. Clin. Neuro.*, 103, 474–485.
- Pérez-González, D. & Malmierca, M.S. (2014) Adaptation in the auditory system: an overview. *Front. Integr. Neurosci.*, 8, 19.
- Picton, T.W., Alain, C., Otten, L., Ritter, W. & Achim, A. (2000) Mismatch negativity: different water in the same river. *Audiol. Neuro-Otol.*, 5, 111– 139.
- Rauschecker, J.P. (1997) Processing of complex sounds in the auditory cortex of cat, monkey, and man. Acta Otolaryngol. Suppl., 532, 34–38.
- Recanzone, G.H. (2008) Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. *J. Neurosci.*, 28, 13184–13193.
- Reyes, A.D. (2011) Synaptic short-term plasticity in auditory cortical circuits. *Hearing Res.*, 279, 60–66.
- Sadagopan, S. & Wang, X. (2009) Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. J. Neurosci., 29, 11192–11202.
- Taaseh, N., Yaron, A. & Nelken, I. (2011) Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One*, 6, e23369.
- Tsodyks, M.V. & Markram, H. (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA*, 94, 719–723.
- Ulanovsky, N., Las, L. & Nelken, I. (2003) Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.*, 6, 391–398.
- Ulanovsky, N., Las, L., Farkas, D. & Nelken, I. (2004) Multiple time scales of adaptation in auditory cortex neurons. J. Neurosci., 24, 10440–10453.
- Varela, J.A., Sen, K., Gibson, J., Fost, J., Abbott, L.F. & Nelson, S.B. (1997) A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J. Neurosci.*, **17**, 7926– 7940.
- Wang, X. & Kadia, S.C. (2001) Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. J. Neurophysiol., 86, 2616–2620.
- Wang, X., Merzenich, M.M., Beitel, R. & Schreiner, C.E. (1995) Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J. Neurophysiol.*, 74, 2685–2706.
- Wang, X., Lu, T., Snider, R.K. & Liang, L. (2005) Sustained firing in auditory cortex evoked by preferred stimuli. *Nature*, 435, 241–346.
- Wehr, M. & Zador, A.M. (2003) Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, 426, 442–446.
- Wehr, M. & Zador, A.M. (2005) Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron*, 47, 437–445.
- Wilson, H.R. & Cowan, J.D. (1972) Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12, 1–24.
- Young, E.D. (2008) Neural representation of spectral and temporal information in speech. *Philos. T. Roy. Soc. B.*, 363, 923–945.