The auditory N1 suppression rebounds as prediction persists over time

Yi-Fang Hsu [a,b,c,*], Jarmo A. Hämäläinen [d], Florian Waszak [a,b]

[a] Université Paris Descartes, Sorbonne Paris Cité, 75006 Paris, France
[b] CNRS (Laboratoire Psychologie de la Perception, UMR 8242), 75006 Paris, France
[c] Department of Educational Psychology and Counselling, National Taiwan Normal University, 10610 Taipei, Taiwan
[d] Jyväskylä Centre for Interdisciplinary Brain Research, Department of Psychology, University of Jyväskylä, 40014 Jyväskylä, Finland

ABSTRACT

The predictive coding model of perception proposes that neuronal responses reflect prediction errors. Repeated as well as predicted stimuli trigger suppressed neuronal responses because they are associated with reduced prediction errors. However, many predictable events in our environment are not isolated but sequential, yet there is little empirical evidence documenting how suppressed neuronal responses reflecting reduced prediction errors change in the course of a predictable sequence of events. Here we conceived an auditory electroencephalography (EEG) experiment where prediction persists over series of four tones to allow for the delineation of the dynamics of the suppressed neuronal responses. It is possible that neuronal responses might decrease for the initial predictable stimuli and stay at the same level across the rest of the sequence, suggesting that they reflect the predictability of the stimuli in terms of mere probability. Alternatively, neuronal responses might decrease for the initial predictable stimuli and gradually recover across the rest of the sequence, suggesting that factors other than mere probability have to be considered in order to account for the way prediction is implemented in the brain. We found that initial presentation of the predictable stimuli was associated with suppression of the auditory N1. Further presentation of the predictable stimuli was associated with a rebound of the component's amplitude. Moreover, such pattern was independent of attention. The findings suggest that auditory N1 suppression reflecting reduced prediction errors is a transient phenomenon that can be modulated by multiple factors.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The brain is believed to specialise in predictive information processing. The predictive coding model of perception postulates that perception entails two distinct neurocomputational components, the top-down propagation of predictions and the bottom-up propagation of prediction errors (Friston, 2005, 2009; Egner et al., 2010; Feldman and Friston, 2010; see Clark, 2013 for a review). While predictions about the causal structure of the world are provided backward to the next lower level, mismatch between predictions and sensory inputs is encoded as prediction errors, which will be communicated forward to the next higher level. Such correction iterates throughout the hierarchy to match predictions and sensory inputs as much as possible to minimise prediction errors in the system.

This model proposes that neuronal responses reflect prediction errors (Friston, 2005, 2009). The account nicely explains repetition suppression, a robust phenomenon of suppressed neuronal responses to repeated stimuli (Henson and Rugg, 2003; Grill-Spector et al., 2006). It is suggested that repetition suppression is due to the repeated stimuli being expected such that there is smaller inconsistency between the anticipated and perceived stimuli, resulting in reduced prediction errors. This idea is supported by recent research showing that repetition suppression was attenuated when the presentation of repeated stimuli was unexpected (functional magnetic resonance imaging (fMRI): Summerfield et al., 2008; Kovács et al., 2012; Grotheer and Kovács, 2014; Mayrhauser et al., 2014; magnetoencephalography (MEG): Todorovic et al. (2011); electroencephalography (EEG): Summerfield et al., 2011). This notion is further corroborated by reports on prediction suppression. In a range of EEG studies, predicted stimuli were found to trigger suppressed sensory event-related potentials (ERPs) such as the visual N1 (Gentsch and Schütz-Bosbach, 2011)

* Correspondence to: Department of Educational Psychology and Counselling, National Taiwan Normal University. 162, Section 1, Heping E. Road, 10610 Taipei, Taiwan. Tel.: +8862 77 34 37 64; fax: +8862 23 41 38 65.
E-mail addresses: yi-fang.hsu@cantab.net (Y.-F. Hsu),
jarmo.a.hamalainen@jyu.fi (J.A. Hämäläinen),
Florian.waszak@parisdescartes.fr (F. Waszak).

and the auditory N1 (Schafer and Marcus, 1973; Schafer et al., 1981; Lange, 2009; SanMiguel et al., 2013; Timm et al., 2013; see Bendixen et al., 2012 for a review).

Note that it is only rarely the case that we predict isolated stimuli in our environment. Usually, we must predict a whole sequence of events. However, there are only a handful of studies that have explicitly investigated how predictions evolve beyond a simple repetition (Auksztulewicz and Friston, 2015; Garrido et al., 2009; Wacongne et al., 2011). This issue, addressed in the current research, is of great interest as an understanding of how sequential prediction works will help to elucidate the mechanism underlying prediction implementation in the brain. Specifically, the question of whether neuronal responses change across the sequence of stimuli will reveal whether suppressed neuronal responses reflecting reduced prediction errors is a static state or a transient phenomenon that can be modulated by multiple factors.

Instead of a series of repetitive tones, the effect of predictions on neuronal responses should be best revealed with a series of tones following abstract rules, because any effect demonstrated cannot be explained by simple inhibitory response of the neuronal populations responsive to the tone frequency. Therefore, here we conceived an auditory EEG experiment where abstract rules were introduced in series of four tones. To control for the effect of attention, which can be conflated with the effect of prediction in the literature (see Summerfield and Egner, 2009 for a review), series of predictable tones were interleaved with series of unpredictable tones. Participants' attention was directed to one of the two interleaved series, creating attended and unattended conditions.

We see two possibilities of how prediction suppression might develop across a predictable sequence. First, neuronal responses might decrease for the initial predictable stimuli and stay at the same level across the rest of the sequence, suggesting that they reflect the predictability of the stimuli in terms of mere probability. The traditional literature on auditory N1 has demonstrated that the auditory N1 showed marked decrease and then stabilised at floor level after repetitive presentation of a tone. This has been suggested to be due to refractory properties of the neuronal populations (i.e., inhibition of neuronal populations with characteristic frequency; see Näätänen and Picton, 1987 for a review). On the other hand, in previous research using a roving oddball sequence, auditory ERPs have been shown to be modulated further by stimulus repetition (Haenschel et al., 2005; Garrido et al., 2009). In these experiments, the initial presentation of an oddball elicited a mismatch negativity (MMN) response which decreased in amplitude to the repetitive presentation of the same stimulus. Importantly, when examining the reconstructed activity originating from the primary auditory cortex, Garrido et al. (2009) found that the third and the fourth repetition of the same stimulus produced an increase in amplitude of the auditory N1 compared to the second stimulus in the sequence. Thus, alternatively, neuronal responses might decrease for the initial predictable stimuli and gradually recover across the rest of the sequence, suggesting that factors other than mere probability have to be considered in order to account for the way prediction is implemented in the brain as indicated by Garrido et al. (2009).

2. Materials and methods

2.1. Participants

Eighteen healthy volunteers (average age 26; nine males; seventeen right-handed) with no history of neurological, psychiatric, or visual/hearing impairments as indicated by self-report participated in the experiment. Participants gave written informed consent and were paid for participation. Ethical approval was granted by the Comité de Protection des Personnes (CPP) Ile de France II.

2.2. Stimuli

Sinusoidal tones with a loudness of 80 phons (i.e., 80 dB for tones of 1000 Hz) were generated using Matlab 2007b (The Mathworks Inc.). The duration of each tone was 50 ms (including 5 ms rise/fall times). The frequency of each tone was within the range of 261.626–493.883 Hz and 2093.000–3951.070 Hz, matching the absolute frequency of two sets of seven natural keys on a modern piano (low-pitched set: C4 D4 E4 F4 G4 A4 B4; high-pitched set: C7 D7 E7 F7 G7 A7 B7).

Within each frequency set, 360 series of four tones were created for the predictable and unpredictable stimulus streams, respectively. In other words, there were respectively 1440 tones in the low-pitched predictable stimulus stream, the high-pitched predictable stimulus stream, the low-pitched unpredictable stimulus stream, and the high-pitched unpredictable stimulus stream. Each participant received only two of these streams (i.e., one predictable and one unpredictable stimulus stream from different frequency sets; see below for details). For the four-tone series in the predictable stimulus stream, the frequency of the first tone was determined by a random selection with equal probability except that it could not be the fourth tone of the previous series. The frequency of the second, the third, and the fourth tone was always one natural key higher than its preceding tone (e.g., C4-D4-E4-F4; C7-D7-E7-F7). This is based on the idea that presenting the frequencies as a scale should lead to the expectation that the frequencies of upcoming tones continue in the same direction (i.e., step inertia) (Lange, 2009). Therefore, while the first tone was a non-predicted baseline, the second, the third, and the fourth tones were predictable in terms of frequency. For the four-tone series in the unpredictable stimulus stream, the frequency of each tone was determined by a random selection with equal probability except for immediate repetition. Therefore, while the first tone was a non-predicted baseline, the second, the third, and the fourth tones were unpredictable in terms of frequency.

Continuous series of tones were created by interleaving the predictable and unpredictable stimulus streams from different frequency sets to allow for the efficient manipulation of attention on the two stimulus streams. That is, a low-pitched predictable stimulus stream was interleaved with a high-pitched unpredictable stream, whereas a high-pitched predictable stimulus stream was interleaved with a low-pitched unpredictable stream. A stimulus onset asynchrony (SOA) of 500 ms was used. E-prime version 2.0 (Psychology Software Tools, Inc., USA) was used for stimulus presentation. The stimulation sequence was randomised individually for each participant and delivered binaurally via headphones (Sennheiser PX200).

2.3. Procedures

A total of twelve blocks of 240 tones were presented, including 120 tones from the predictable stimulus stream and 120 tones from the unpredictable stimulus stream in each block. To counterbalance the effect of frequency set, half of the participants were given a low-pitched predictable stimulus stream interleaved with a high-pitched unpredictable stream, whereas half of the participants were given a high-pitched predictable stimulus stream interleaved with a low-pitched unpredictable stream. Orthogonal to this, half of the participants were instructed to pay attention to the low-pitched stimulus stream in the first six blocks and then the high-pitched stimulus stream in the remaining six blocks, whereas half of the participants were instructed to pay attention to the high-pitched stimulus stream in the first six blocks and then the low-pitched stimulus stream in the remaining six blocks. In this
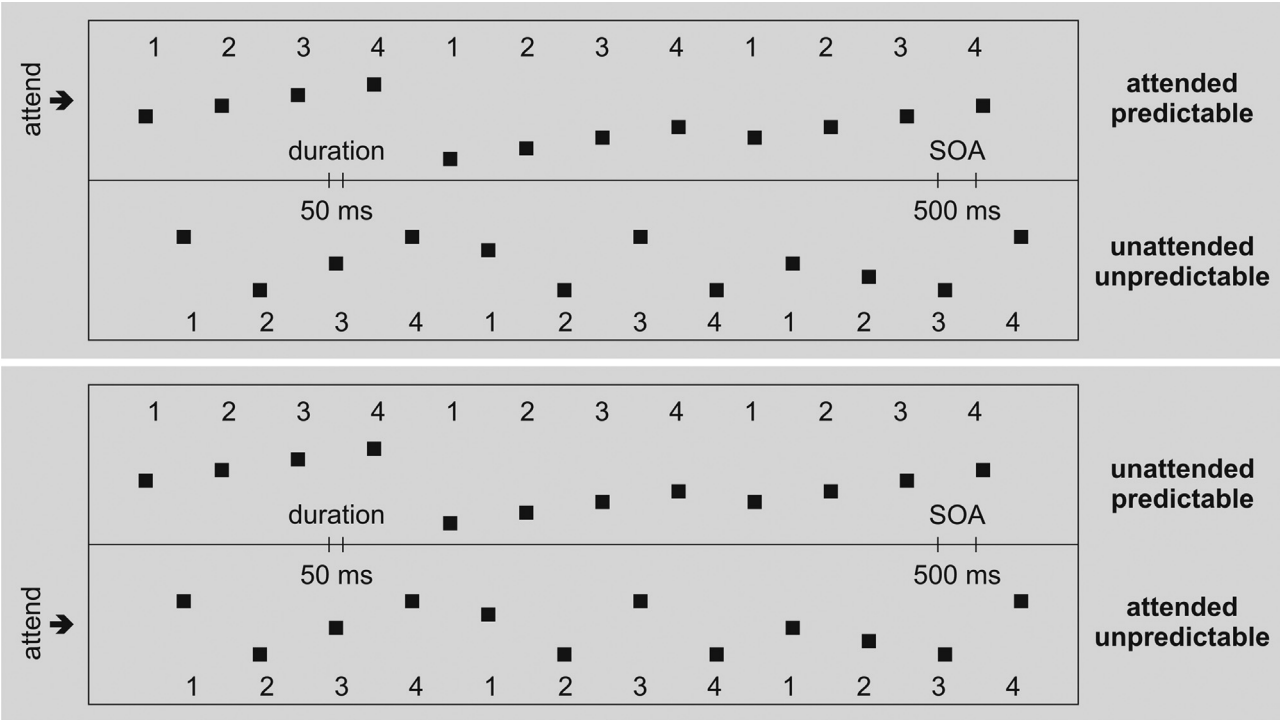
Fig. 1. Schematic illustration of the paradigm, in which the predictable and unpredictable stimulus stream are interleaved.

case, participants' attention was directed to the predictable and unpredictable stimulus streams in 50% of the blocks, respectively (Fig. 1). Note that information concerning stimulus predictability was not mentioned in the instruction. Throughout the experiment, 10% of the tones were attenuated by 20 dB in the attended stimulus stream to serve as targets and 10% of the tones were attenuated by 20 dB in the unattended stimulus stream to serve as distractors. Participants were required to press a key when they detected a softer tone in the attended stimulus stream which randomly occurred 12 times in each block, ignoring the stimulation in the unattended stimulus stream. The experiment took around 24 min.

2.4. Data recording and analysis

2.4.1. EEG recording and pre-processing

EEG was recorded with 64 active electrodes (actiCAP, Brain Products GmbH, Germany) conforming to the international 10-10 system. The sampling rate was 500 Hz. No online/offline filter was used. The Cz served as the reference electrode online. The data was recomputed to average reference offline. Target stimuli, the first stimuli following target stimuli (where there can be interference from movement-related potentials), and distractor stimuli were removed. Epochs extended from -100 ms to 500 ms relative to stimulus onset, using a 100 ms pre-stimulus baseline. Ocular artefact correction was conducted with independent component

analysis in EEGlab (Delorme and Makeig, 2004). Epochs containing voltage deviations exceeding ± 100 μV relative to baseline at any of the electrodes were rejected. The trial numbers after artefact rejection in each condition are listed in Table 1.

2.4.2. ERP analysis

ERP analysis was based on a temporal principal component analysis (PCA) in SPSS 20. The temporal PCA statistically decomposes the ERP waveforms into constituent building blocks, which affords objective data-driven ERP component measures when compared to the conventional peak-picking methods (Kayser and Tenke, 2003, 2006; Dien and Frishkoff, 2005; Dien, 2012). Moreover, it is not susceptible to the influences of high-frequency noises and low-frequency drifts in the data as the conventional peak-picking methods (Luck, 2005). Covariance matrix and Promax rotation were used. All components accounting for a total of 99% of the variance (maximum iterations for convergence=500) were included in the rotation (Promax Kappa=4). The temporal decomposition provided a set of time-variant component loadings reflecting the contribution of each temporal component to the voltage at each time point and a set of time-invariant component scores (calculated using Barlett method) representing the contribution of each data point (i.e., the contribution of different electrodes, participants, and conditions) to the principal component which can be subject to inferential statistics (Van Boxtel, 1998). The component corresponding to the auditory N1 (i.e.,

Table 1
Mean and range of trial numbers after artefact rejection in each condition.

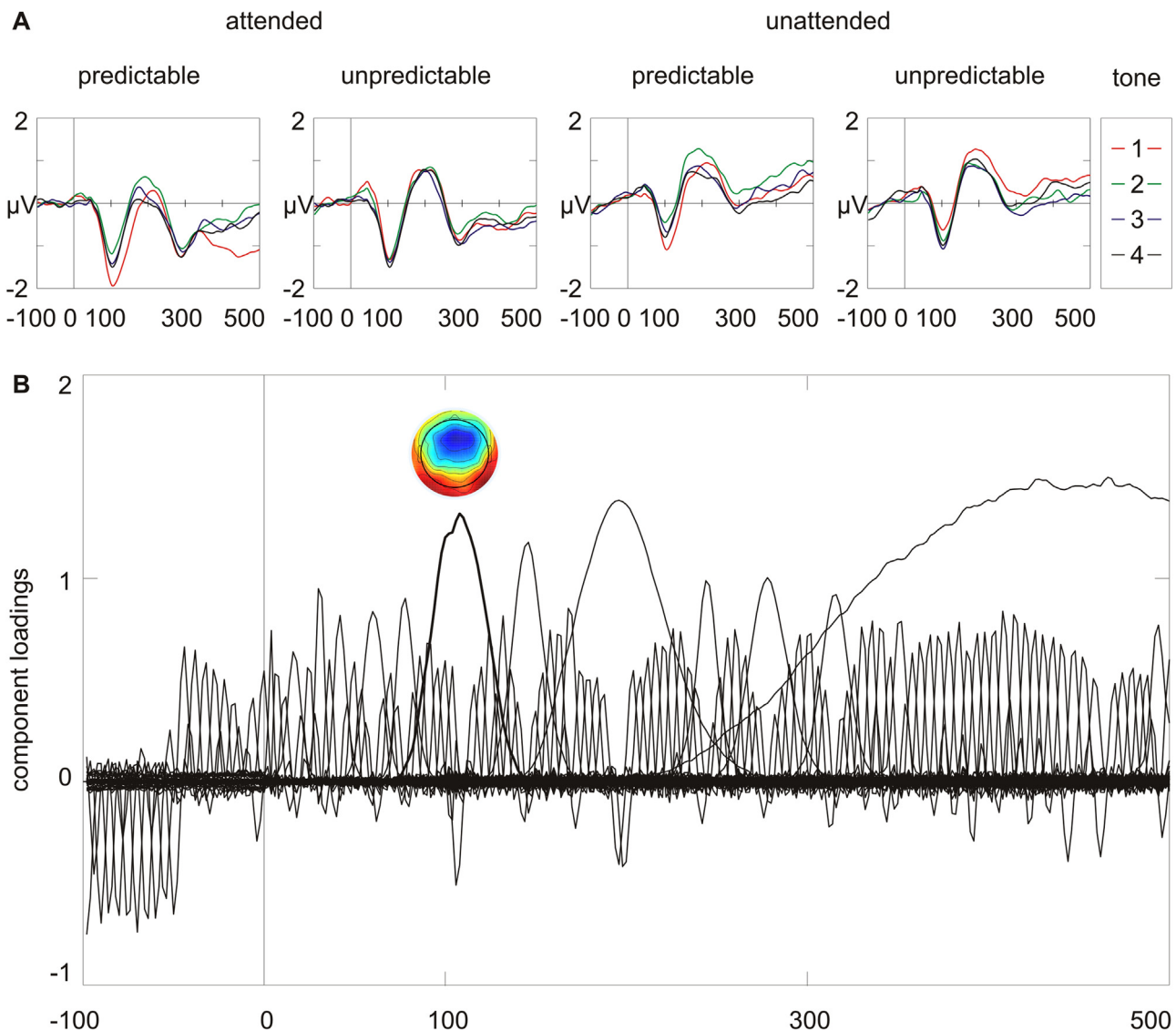| Mean (range)	Tone 1	Tone 2	Tone 3	Tone 4
Attended predictable	155.33 (114–167)	153.17 (117–163)	154.17 (114–163)	153.83 (119–165)
Attended unpredictable	152.06 (124–165)	150.44 (123–162)	152.00 (122–166)	154.39 (124–166)
Unattended predictable	137.72 (116–146)	137.33 (108–151)	135.00 (102–148)	138.50 (103–153)
Unattended unpredictable	139.28 (106–150)	137.56 (109–145)	139.33 (116–148)	139.33 (108–148)

Fig. 2. (A) Grand average ERPs on Fz lowpass filtered at 20 Hz for visual presentation purposes. (B) Component loadings of 146 components in the temporal PCA. The component corresponding to the auditory N1 in the grand average ERPs is marked with a thick line with its topographical distribution plotted on top.

principal component 3 accounting for 7.00% of the variance) was identified on the basis of the component loading latency and the component score topography (see Fig. 2B). The component scores were averaged across three electrodes showing the largest component score across all conditions independent of experimental manipulation (i.e., Fz, F2, F4). The advantage of averaging three maximum electrodes was twofold. First, it increased the signal-to-noise ratio of the component. Second, it avoided the problems inherited in the analysis of predefined areas that takes an average of multiple electrodes over predefined regions, which might not correspond to the true topography in the experiment. Moreover, these electrodes representing the locations of maximum signals for the auditory N1 corresponded to those reported in the literature (e.g., Näätänen and Picton, 1987). Therefore, the averages of these electrodes can be considered as objective representatives of the component to serve as inputs for a 2 (attention: attended/unattended)x2 (prediction: predictable/unpredictable)x4 (position: tone 1/2/3/4) repeated measures analysis of variance (ANOVA). A Greenhouse-Geisser correction was applied when sphericity was violated.

3. Results

3.1. Behavioural results

Overall, participants' behavioural performance in the target detection task was close to ceiling (hit rate: mean$=0.84$, SD$=0.10$; false alarm rate: mean< 0.01, SD< 0.01; RT: mean$=531.36$, SD$=57.85$). There was no difference between participants' behavioural performance when they attended to predictable and unpredictable stimulus stream (hit rate: $t(17)=1.83$, $p=0.08$; false alarm rate: $t(17)=1.69$, $p=0.11$; RT: $t(17)=-1.75$, $p=0.10$), suggesting that task difficulty was equivalent across blocks.

3.2. ERP results

Fig. 2A shows the grand average ERPs on Fz lowpass filtered at 20 Hz for visual presentation purposes. Fig. 2B shows the component loadings of 146 components in the temporal PCA. The component corresponding to the auditory N1 in the grand average ERPs was identified on the basis of the component loading latency (i.e., 100 ms) and the component score topography (i.e., fronto-central negativity).
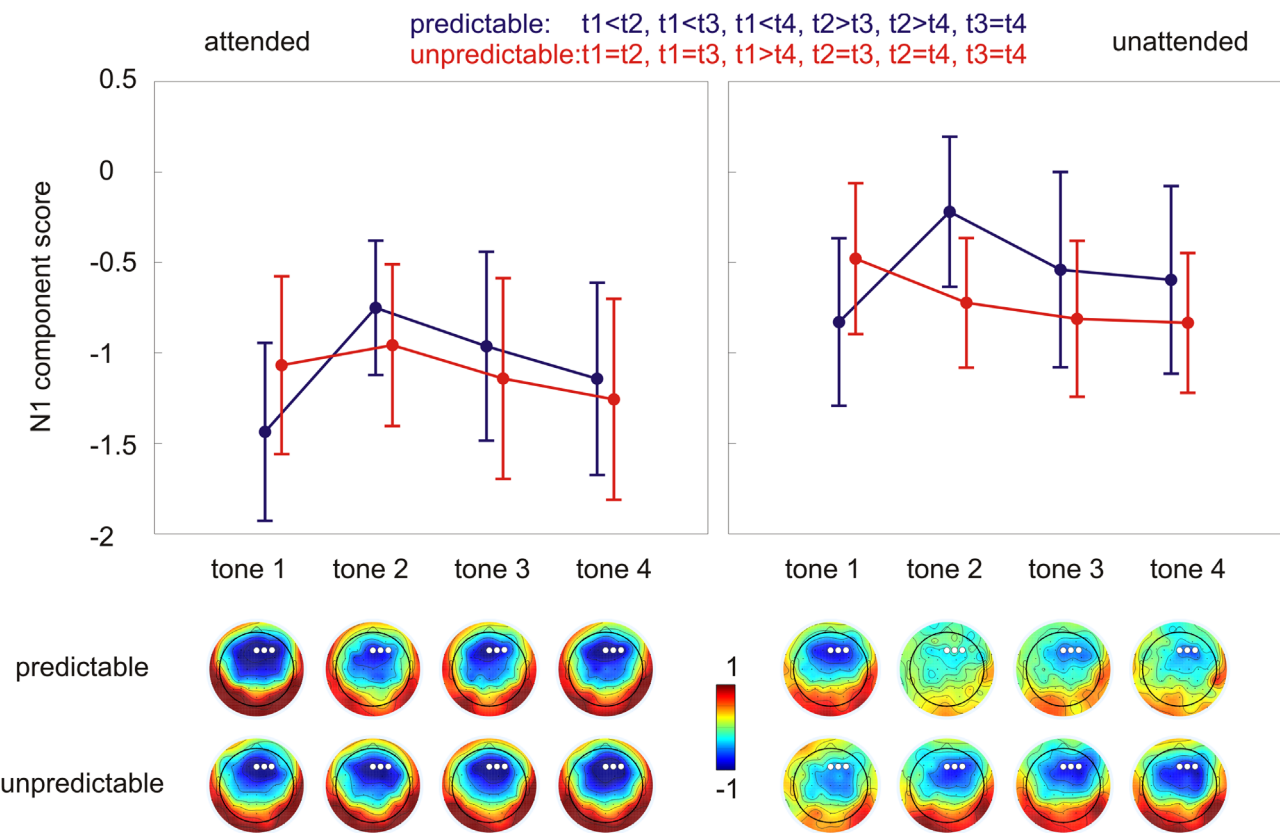
Fig. 3. The auditory N1 component score averaged across three electrodes showing the largest response (i.e., Fz, F2, F4; marked as white dots on the component score topographic maps) in each condition. Error bars depict one standard deviation of the mean.

Fig. 3 shows the auditory N1 component score averaged across three electrodes showing the largest response in each condition. There was no three-way interaction between attention, prediction, and position ($F_{(3,51)}=0.40$, $p=0.75$, partial $\eta^2=0.02$). The main effect of attention was significant ($F_{(1,17)}=21.20$, $p<0.001$, partial $\eta^2=0.56$), but it did not interact with any other variables (attention x prediction: $F_{(1,17)}=0.33$, $p=0.57$, partial $\eta^2=0.02$; attention x position: $F_{(3,51)}=1.10$, $p=0.36$, partial $\eta^2=0.06$). On the other hand, there was a significant two-way interaction between prediction and position ($F_{(3,51)}=12.81$, $p<0.001$, partial $\eta^2=0.43$).

Post hoc comparisons showed that the significant main effect of position was found in predictable conditions (attended predictable condition: $F_{(3,51)}=9.63$, $p<0.001$, partial $\eta^2=0.36$; unattended predictable condition: $F_{(3,51)}=5.93$, $p\leq0.001$, partial $\eta^2=0.26$) but not in unpredictable condition (attended unpredictable condition: $F_{(2.11,35.88)}=1.63$, $p=0.21$, partial $\eta^2=0.09$; unattended unpredictable condition: $F_{(3,51)}=2.82$, $p=0.05$, partial $\eta^2=0.14$). Therefore, the four conditions were collapsed into predictable and unpredictable conditions in the pairwise comparisons as shown in Table 2. In the predictable condition, pairwise comparison showed that tone 1 response was significantly different from tone 2 response, tone 3 response, and tone 4 response. In addition, tone 2 response was significantly different from tone 3 response and 4 response, while tone 3 response did not differ from tone 4 response. In the unpredictable condition, pairwise comparisons showed that tone 1 response was significantly different from tone 4 response. No other contrast was significant.

4. Discussions

The predictive coding model of perception proposes that neuronal responses reflect prediction errors (Friston, 2005, 2009).

Table 2
Pairwise comparisons of the auditory N1 amplitudes. Attended and unattended conditions have been collapsed.

Contrast	Mean difference	Standard error	p-value	95% confidence interval
Predictable				
Tone 1 vs. tone 2	−0.65	0.10	<0.001	−0.86 to −0.43
Tone 1 vs. tone 3	−0.38	0.11	<0.004	−0.61 to −0.15
Tone 1 vs. tone 4	−0.26	0.10	<0.019	−0.48 to −0.05
Tone 2 vs. tone 3	0.27	0.11	<0.026	0.04–0.49
Tone 2 vs. tone 4	0.38	0.11	<0.004	0.15–0.62
Tone 3 vs. tone 4	0.11	0.11	<0.325	−0.12–0.34
Unpredictable				
Tone 1 vs. tone 2	0.07	0.08	<0.433	−0.11–0.24
Tone 1 vs. tone 3	0.20	0.10	<0.051	0.00–0.41
Tone 1 vs. tone 4	0.27	0.09	<0.010	0.08–0.47
Tone 2 vs. tone 3	0.14	0.08	<0.118	−0.04–0.31
Tone 2 vs. tone 4	0.21	0.10	<0.064	−0.01–0.42
Tone 3 vs. tone 4	−0.04	0.10	<0.712	−0.24–0.17

Repeated as well as predicted stimuli triggered suppressed neuronal responses because they are associated with reduced prediction errors (repetition suppression: Summerfield et al., 2008, 2011; Todorovic et al., 2011; Kovács et al., 2012; Grotheer and Kovács, 2014; Mayrhauser et al., 2014; prediction suppression: Schafer and Marcus, 1973; Schafer et al., 1981; Lange, 2009; Gentsch and Schütz-Bosbach, 2011; SanMiguel et al., 2013; Timm et al., 2013). However, many predictable events in our environment are not isolated but sequential, such as the tune of a melody, yet there is little empirical evidence documenting how neuronal responses reflecting prediction errors change in the course of a predictable sequence of events. Here we found a suppression followed by a rebound on the amplitude of the auditory N1 in a

sequence of equally predictable events.

The overall pattern of results seems to disagree with the notion that the auditory N1 reflects the predictability of the event in terms of mere probability, which was implied in previous research showing that the auditory N1 decreased as prediction errors decreased (Schafer and Marcus, 1973; Schafer et al., 1981; Lange, 2009; SanMiguel et al., 2013; Timm et al., 2013; see Bendixen et al., 2012 for a review). In particular, it is in contrast with one line of previous research reporting that the auditory N1 showed marked decrease and then stabilised at floor level after repetitive presentation of a sound (see Näätänen and Picton, 1987 for a review), with one interpretation of the effect being related to repetition positivity, a frontocentral slow positive wave at 50–250 ms increasing with the number of stimulus repetition (Baldeweg et al., 2004; Haenschel et al., 2005; Costa-Faidella et al., 2011). This repetition positivity has been suggested to reflect an increase in the predictability of the repeated tone due to better encoding of the stimulus features (Baldeweg et al., 2004; Haenschel et al., 2005). On the contrary, our results are in line with the finding of Garrido et al. (2009) where, in a sequence of repetitive oddball stimuli, the reconstructed activity from A1 showed a decrease from the first tone to the second tone and then an increase for the following tones. This effect was maximal before 100 ms matching rather closely with the time window of the effect reported above. The current research corroborates and extends their finding in that it adopts proper refractoriness controls, as we investigated the dynamics of the suppressed neuronal responses without manipulating stimulus repetition but stimulus regularity. Specifically, the four-tone series in the predictable stimulus stream was presented as a scale, so that the frequency of the first tone can predict the frequency of the second, the third, and the fourth tones. Our design therefore shows how the auditory N1 changes in a sequence of predictable events excluding the effect of refractoriness (Budd et al., 1998).

There are at least three possible explanations for the result pattern observed in the current research, all suggesting that factors other than mere probability have to be considered in order to account for the way prediction is implemented in the brain. First, the auditory N1 suppression might be modulated by the narrowing bandwidth of the prediction tuning curve over time. As the brain constantly attempts to minimise prediction errors in the system, the increased confidence in prediction might result in sparser representation of predictions (cf. sharpening model for repetition effect: Desimone, 1996; Wiggs and Martin, 1998; see Grill-Spector et al., 2006 for a review). That is, the increased confidence in prediction might sharpen the neuronal responses around the preferred prediction, narrowing the bandwidth of the prediction tuning curve. Given that our prediction of frequency is rarely absolutely precise, this in turn paradoxically elicits an increase in prediction errors. Therefore, while earlier predictions are associated with a suppression, later predictions are associated with a rebound. Second, the auditory N1 suppression might be weighted by a neuronal learning function of escalating sensitisation. While the second, the third, and the fourth tones are associated with equal amount of suppression, they can be weighted by a mounting function reflecting the build-up of representations. For example, a study using children as participants found that a vertex neuronal response can increase its amplitude along trains of auditory stimulation, probably reflecting the development of the representation for the sounds (Karhu et al., 1997). The third possibility is that the rebound of the auditory N1 might be due to increased expectation for the onset of the next four-tone series. These three possibilities raise testable hypotheses for future research. Nevertheless, the result pattern suggests that auditory N1 suppression reflecting reduced prediction errors is a transient phenomenon that can be modulated by multiple factors. Whether the result pattern may change over time, interacting with the next level of prediction in the hierarchy, remains an open question requiring systematic investigations.

We would also like to point out that the auditory N1 to the first tone in the predictable stimulus stream could be partly overlapping with a MMN response. For example, Wacongne et al. (2011) presented participants with blocks of stimuli consisting of rare XXXXX sound pattern (i.e., five identical tones) and frequent XXXXY sound pattern (i.e., four identical tones followed by a different tone), the latter of which is analogous to the stimuli used in the current research in that one four-tone series forms a linear pattern that is then broken by the first tone of the next four-tone series. Wacongne et al. (2011) found that, although XXXXY was an expected sound pattern, the Y stimulus triggered a small MMN response. This suggests that, in the current research, part of the neuronal enhancement to the first tone in the auditory N1 time window could be attributed to a MMN response.

Importantly, the overall pattern of results (i.e., the suppression followed by the rebound) was observed in the predictable stimulus stream, not in the unpredictable stimulus stream. The interaction between predictability (predictable/unpredictable) and position (tone 1/2/3/4) therefore suggest that our finding is not simply a reflection of the passage of time. This can be related to previous research on continuous speech segmentation showing that the difference between auditory N1 elicited by initial syllables and medial syllables of nonsense words was larger after than before training (Sanders et al., 2002), suggesting that the auditory N1 effect indexes the involvement of predictive processes.

Lastly, our result pattern was independent of attention. While we found the auditory N1 to be enhanced under attended conditions, consistent with previous reports on early effects of attention on brain responses (Näätänen and Picton, 1987; Chennu et al., 2013), the effects of attention (i.e., an enhancement) and prediction (i.e., a suppression followed by a rebound) were clearly independent of each other. The independency of the two effects is in line with the findings that repeated sounds are associated with suppressed auditory N1 regardless of attention allocation (Haenschel et al., 2005; Hsu et al., 2014a). It is also consistent with reports on prediction suppression that the suppressed auditory N1 for self-initiated sounds was not modulated by attention allocation (Timm et al., 2013). Taken together, these studies seem to indicate that the prediction effect on the auditory N1 is rather automatic. On the other hand, how do these results reconcile with previous research suggesting prediction effect to be attention-dependent? One possibility is that this difference is related to how straightforward predictions can be formed in the particular experiment. In the current as well as previous research reporting the prediction mechanism to be independent of attention, participants' prediction was based on pre-existing rules, such as step inertia, repetition tendency, and action-effect binding where voluntary actions were believed to trigger sensory effects in the experiments (e.g., Gentsch and Schütz-Bosbach, 2011; Kovács et al., 2012). In contrast, in previous research suggesting the prediction effect to be attention-dependent, participants' prediction was based on arbitrary regularities introduced in the experiments (e.g., Larsson and Smith, 2010; Hsu et al., 2014b). Future research is needed to investigate whether the attention-dependency of the prediction effect is related to how prediction is manipulated.

Conflict of interest

The authors declare no competing financial interests.

Acknowledgement

References

Auksztulewicz, R., Friston, K., 2015. Attentional enhancement of auditory mismatch responses: a DCM/MEG study. Cereb. Cortex . http://dx.doi.org/10.1093/cercor/bhu323.

Baldeweg, T., Klugman, A., Gruzelier, J., Hirsch, S.R., 2004. Mismatch negativity potentials and cognitive impairment in schizophrenia. Schizophr. Res. 69, 203–217.

Bendixen, A., SanMiguel, I., Schröger, E., 2012. Early electrophysiological indicators for predictive processing in audition: a review. Int. J. Psychophysiol. 83, 120–131.

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 (1), 51–68.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñes, A., Owen, A.M., Bekinschtein, T.A., 2013. Expectation and attention in hierarchical auditory prediction. J. Neurosci. 33, 11194–11205.

Clark, A., 2013. Whatever next? predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36 (3), 181–204.

Costa-Faidella, J., Grimm, S., Slabu, L., Díaz-Santaella, F., Escera, C., 2011. Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. Psychophysiology 48, 774–783.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21.

Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. Proceedings of the National Academy of Sciences 93, 13494–13499.

Dien, J., Frishkoff, G.A., 2005. Principal components analysis of event-related potential datasets. In: Handy, T. (Ed.), Event-related potentials: A Methods Handbook. MIT Press, Cambridge, MA, pp. 189–208.

Dien, J., 2012. Applying principal components analysis to event-related potentials: a tutorial. Dev. Neuropsychol. 37 (6), 497–517.

Egner, T., Monti, J.M., Summerfield, C., 2010. Expectation and surprise determine neural population responses in the ventral visual stream. J. Neurosci. 30 (49), 16601–16608.

Feldman, H., Friston, K.J., 2010. Attention, uncertainty, and free-energy. Frontiers in human neuroscience 4, 215.

Friston, K.J., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B 360, 815–836.

Friston, K.J., 2009. The free-energy principle: a rough guide to the brain? Trends Cognit. Sci. 13 (7), 293–301.

Garrido, M.I., Kilner, J.M., Kiebel, S.J., Stephan, K.E., Baldeweg, T., Friston, K.J., 2009. Repetition suppression and plasticity in the human brain. NeuroImage 48, 269–279.

Gentsch, A., Schütz-Bosbach, S., 2011. I did it: unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. J. Cognit. Neurosci. 23 (12), 3817–3828.

Grill-Spector, K., Henson, R.N.A., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cognit. Sci. 10 (1), 14–23.

Grotheer, M., Kovács, G., 2014. Repetition probability effects depend on prior experiences. J. Neurosci. 34 (19), 6640–6646.

Haenschel, C., Vernon, D.J., Dwivedi, P., Gruzelier, J.H., Baldeweg, T., 2005. Event-related brain potential correlates of human auditory sensory memory-trace formation. J. Neurosci. 25, 10494–10501.

Henson, R.N.A., Rugg, M.D., 2003. Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia 41, 263–270.

Hsu, Y.-F., Hämäläinen, J.A., Waszak, F., 2014a. Repetition suppression comprises both attention-independent and attention-dependent processes. NeuroImage 98, 168–175.

Hsu, Y.-F., Hämäläinen, J.A., Waszak, F., 2014b. Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. Frontiers in human neuroscience 8, 152.

Karhu, J., Herrgård, E., Pääkkönen, A., Luoma, L., Airaksinen, E., Partanen, J., 1997. Dual cerebral processing of elementary auditory input in children. Neuroreport 8 (6), 1327–1330.

Kayser, J., Tenke, C.E., 2003. Optimizing PCA methodology for ERP component identification and measurement: theoretical rationale and empirical evaluation. Clin. Neurophysiol. 114, 2307–2325.

Kayser, J., Tenke, C.E., 2006. Consensus on PCA for ERP data, and sensibility of unrestricted solutions. Clin. Neurophysiol. 117, 695–707.

Kovács, G., Iffland, L., Vidnyánszky, Z., Greenlee, M.W., 2012. Stimulus repetition probability effects on repetition suppression are position invariant for faces. Neuroimage 60, 2128–2135.

Lange, K., 2009. Brain correlates of early auditory processing are attenuated by expectations for time and pitch. Brain Cogn. 69 (1), 127–137.

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

Mayrhauser, L., Bergmann, J., Crone, J., Kronbichler, M., 2014. Neural repetition suppression: evidence for perceptual expectation in object-selective regions. Frontiers in human neuroscience 8, 225.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology 24, 375–425.

Sanders, L.D., Newport, E.L., Neville, H.J., 2002. Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. Nat. Neurosci. 5 (7), 700–703.

SanMiguel, I., Todd, J., Schröger, E., 2013. Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. Psychophysiology 50 (4), 334–343.

Schafer, E.W., Amochaev, A., Russell, M.J., 1981. Knowledge of stimulus timing attenuates human evoked cortical potentials. Electroencephalogr. Clin. Neurophysiol. 52 (1), 9–17.

Schafer, E.W., Marcus, M.M., 1973. Self-stimulation alters human sensory brain responses. Science 181 (4095), 175–177.

Summerfield, C., Egner, T., 2009. Expectation (and attention) in visual cognition. Trends Cognit. Sci. 13 (9), 403–409.

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. 11 (9), 1004–1006.

Summerfield, C., Wyart, V., Johnen, V.M., de Gardelle, V., 2011. Human scalp electroencephalography reveals that repetition suppression varies with expectation. Frontiers in human neuroscience 5, 67.

Timm, J., SanMiguel, I., Saupe, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. BMC Neuroscience 14 (1), 2.

Todorovic, A., van Ede, F., Maris, E., de Lange, F.P., 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J. Neurosci. 31 (25), 9118–9123.

Van Boxtel, G.J.M., 1998. Computational and statistical methods for analyzing event-related potential data. Behav. Res. Methods Instrum. Comput. 30 (1), 87–102.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., Dehaene, S., 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proceedings of the National Academy of Sciences 108, 20754–20759.

Wiggs, C.L., Martin, A., 1998. Properties and mechanisms of perceptual priming. Curr. Opin. Neurobiol. 8, 227–233.