

Serveur Académique Lausannois SERVAL [serval.unil.ch](http://serval.unil.ch)

## Publisher's version PDF

Faculty of Biology and Medicine Publication

Originally published at:

**Title:** Experience-based Auditory Predictions Modulate Brain Activity to Silence as do Real Sounds.

**Authors:** Chouiter L, Tzovara A, Dieguez S, Annoni JM, Magezi D, De Lucia M, Spierer L

**Journal:** Journal of cognitive neuroscience

**Year:** 2015 Oct

**Volume:** 27

**Issue:** 10

**Pages:** 1968-80

**DOI:** [10.1162/jocn\\_a\\_00835](https://doi.org/10.1162/jocn_a_00835)

© (year) Massachusetts Institute of Technology

# Experience-based Auditory Predictions Modulate Brain Activity to Silence as Do Real Sounds

Leila Chouiter<sup>1\*</sup>, Athina Tzovara<sup>2,3,4\*</sup>, Sebastian Dieguez<sup>1</sup>, Jean-Marie Annoni<sup>1</sup>, David Magezi<sup>1</sup>, Marzia De Lucia<sup>2,3\*\*</sup>, and Lucas Spierer<sup>1\*\*</sup>

## Abstract

■ Interactions between stimuli's acoustic features and experience-based internal models of the environment enable listeners to compensate for the disruptions in auditory streams that are regularly encountered in noisy environments. However, whether auditory gaps are filled in predictively or restored a posteriori remains unclear. The current lack of positive statistical evidence that internal models can actually shape brain activity as would real sounds precludes accepting predictive accounts of filling-in phenomenon. We investigated the neurophysiological effects of internal models by testing whether single-trial electrophysiological responses to omitted sounds in a rule-based sequence of tones with varying pitch could be decoded from the responses to real sounds and by analyzing the ERPs to the omissions with data-driven electrical neuroimaging methods.

The decoding of the brain responses to different expected, but omitted, tones in both passive and active listening conditions was above chance based on the responses to the real sound in active listening conditions. Topographic ERP analyses and electrical source estimations revealed that, in the absence of any stimulation, experience-based internal models elicit an electrophysiological activity different from noise and that the temporal dynamics of this activity depend on attention. We further found that the expected change in pitch direction of omitted tones modulated the activity of left posterior temporal areas 140–200 msec after the onset of omissions. Collectively, our results indicate that, even in the absence of any stimulation, internal models modulate brain activity as do real sounds, indicating that auditory filling in can be accounted for by predictive activity. ■

## INTRODUCTION

Ample evidence indicates that the auditory system automatically extracts regularities from the acoustic environment and builds up, on this basis, internal models enabling one to generate prediction about forthcoming information (Yabe et al., 1998). Although missing information in auditory streams has initially been accounted for as being compensated retrospectively based on the information surrounding the gaps (Bregman, 1994), recent evidence suggests that gaps might rather be filled in based on predictive activity (e.g., Bendixen, Schroger, & Winkler, 2009; Dubnov, 2008; Schroger, Bendixen, Trujillo-Barreto, & Roeber, 2007; Baldeweg, 2006; Zanto, Snyder, & Large, 2006; Winkler, Karmos, & Näätänen, 1996). According to prospective accounts of auditory filling in, the brain activity driven by internal models would not only help compensate for disruptions in auditory streams but also optimize the processing of incoming auditory information and facilitate the detection of novel auditory events. This assumption, however, implies that internal models actually contain consistent information

on environmental regularities and could generate patterns of brain activity matching those elicited by real physical stimulation. To test this premise of prospective accounts, this study investigated whether and how predictions from internal models actually shape brain activity. The current experiment is not designed to discriminate between the existence of either retrospective or predictive mechanisms, nor is it designed to measure the relative strength of contribution of each mechanism. Rather, our experiment and analyses are designed to investigate direct statistical evidence in support of the predictive filling-in phenomenon.

Most current knowledge on how internal models build up and impact perceptual processes comes from electrophysiological “oddball” paradigms in which rare, deviant sounds are presented in a repetitive sequence of identical sounds. As compared with the brain responses to the frequent, predictable sounds, responses to deviant sounds elicit specific brain responses depending on the difference between expected versus actual stimuli (e.g., mismatch negativity component in electrophysiological oddball studies; Schroger et al., 2007; see Bendixen, SanMiguel, & Schroger, 2012, or Näätänen, Paavilainen, Rinne, & Alho, 2007, for review; Sussman, 2007; Winkler et al., 1996). Extreme forms of oddball in which a sound is omitted—rather than being modified—in an otherwise

<sup>1</sup>University of Fribourg, <sup>2</sup>University of Lausanne, <sup>3</sup>University Hospital of Lausanne, <sup>4</sup>University of Zürich

\*Equal contribution.

\*\*Equal contribution.

predictable sequence of sounds allow a direct assessment of the influence of internal models on neural activity because neurophysiological responses to the omissions are not contaminated by the presentation of a real sound (Bendixen et al., 2009; Raij, McEvoy, Makela, & Hari, 1997).

Omission studies so far have suggested a high degree of similarity between the brain responses to expected but omitted sounds and real sounds (Sanmiguel, Saupe, & Schroger, 2013; Wacongne et al., 2011; Bendixen et al., 2009; Janata, 2001; Raij et al., 1997) and reported that attention might modulate brain responses to omissions by changing their latencies or whether cross-modal information is encoded in the internal models of forthcoming auditory information (Bendixen et al., 2012, for review). Activity within the auditory system resembling the activity elicited by actual stimulation has also been reported in literature that did not focus on omissions within a sound sequence but during silence gaps in familiar musical pieces (Kraemer, Macrae, Green, & Kelley, 2005), violations in learned motor–auditory or visuo-auditory coupling (Stekelenburg & Vroomen, 2015; SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schroger, 2013), or attention to auditory events (Voisin, Bidet-Caulet, Bertrand, & Fonlupt, 2006).

However, the literature so far suffers two main limitations. First, although “similarities” between the brain activity during expected but missing sounds and real sounds have been suggested, to our knowledge, there is no direct positive statistical evidence that the brain activity during auditory gaps actually matches that in response to real sounds. This limitation pertains to the utilization of classical null hypothesis testing approach, which cannot demonstrate the absence of a difference between two conditions (e.g., Rouder, Speckman, Sun, Morey, & Iverson, 2009). For instance, Bendixen et al. (2009) showed that ERPs from predictable tones and tone omissions were not significantly different from each other [ $t(13) = -0.517, p = .614$ ] and concluded that “the auditory system preactivates the neural circuits for expected input, using sequential predictions to specifically prepare for future acoustic events.”

In addition, the precise spatio-temporal dynamics of brain responses to expected but omitted sounds remains unclear. Previous ERP studies on auditory internal models limited their investigation to some electrode sites and periods of interest selected a priori based on where and when typical auditory evoked potential components typically manifest (e.g., MMN at 150 msec on fronto-central recording sites in Shinozaki et al., 2003). This approach could be appropriate if brain responses to omitted sounds are similar to responses to real sounds, but this assumption lacked direct empirical support. Other functional neuroimaging approaches on brain responses to omissions or gaps in auditory streams provided compelling spatial information on the brain activity driven by internal models but with low temporal resolution (Kraemer et al., 2005).

To resolve these issues, we analyzed electrical neuroimaging responses to omissions presented in a rule-based sequence of three tones differing in pitch. Within the sequence, the precise onset and direction of the change in pitch of the forthcoming tones were mostly predictable. To investigate the effects of attention on the building up and functional consequences of internal models, the sound sequence was presented either passively or actively in two different groups of participants.

To determine if the brain activity elicited by the predictions from internal models actually matched those evoked by real stimuli, we examined whether the brain responses to the omission could be decoded based on the responses to the real sounds using a single-trial EEG topographic classification procedure. To characterize the precise spatio-temporal brain dynamics of the brain activity during the omissions, we conducted global topographic analyses of the ERPs to the omitted tones with data-driven time-frame-wise randomization statistics. First, we examined whether and when there was a brain response during the expected but omitted tones using a topographic consistency test (TCT). Second, we conducted a topographic and distributed electrical source estimation analysis of the ERP to the omissions to determine whether, when, and where the brain networks responding to the omissions differed depending on their expected pitch.

## METHODS

### Participants

Twenty-six healthy right-handed male volunteers participated in the study (laterality was assessed using the Edinburgh questionnaire by Oldfield, 1971): 13 were ascribed to the passive listening group (aged 19–43 years, mean  $\pm$  *SD* = 25.9  $\pm$  6.1 years), and 13 were ascribed to the active listening group (aged 22–32 years, mean  $\pm$  *SD* = 25.9  $\pm$  3.3 years). Before the experiment, each participant completed a questionnaire assessing general health and musical training (playing an instrument, singing/music lessons).

No participant had a history of neurological or psychiatric illness, and all reported normal hearing. None were musicians. Each participant provided written informed consent to participate in the study. All procedures were approved by the local ethics committee.

### Stimuli

Stimuli were 70-msec pure tones generated using Adobe Audition 2.0 and presented via etymotic ERP4 insert earphones at a level judged the most comfortable by the participants. All chosen presentation levels were between 80 and 90 dB SPL. The frequency of the tones was 700 Hz for the low-pitch (L), 900 Hz for the medium-pitch (M),

and 1100 Hz for the high-pitch (H) conditions, each of which is easily discriminable from the others.

## Procedure and Tasks

Participants were seated in front of an LCD display screen in a dark, sound-isolated, and electrically shielded booth and participated either in the active or passive listening condition.

In the passive listening condition, they were instructed to watch a subtitled silent movie and to ignore the auditory stimuli. Ten blocks of 1000 stimuli each (tones + omissions) were presented during the passive listening session.

Each block contained a pseudorandom sequence of tones (i.e., random order but without repetition) with an average of 900 trials (ca 300 H, 300 M, and 300 L) across blocks of presentation, plus an average of 100 "omissions" (O) pseudorandomly (i.e., without repetition) interleaved in the sequence (Figure 1A). In total, approximately the same numbers of each possible combination between the H, M, and L tones (H-L, H-M, M-H, M-L, L-M, and L-H) and of each possible combination between a tone and an omission (M-O, H-O, and L-O) were presented during the whole listening session. The total numbers of presentations for each tone type were 2991 (H), 3035 (M), 2964 (L), and 1000 (O) across blocks in the passive condition (there was not exactly the same number of each tone presented during the experimental session because the sound sequences were generated using a Markov chain stochastic statistical model).

In the active listening condition, the sequences were the same as in the passive condition, except that participants were instructed to listen to the sounds while visually fixating a central cross on a black background. Before the experiment, the three tones (H, M, and L) were presented to the participants, and one of them was designed as the "target sound" (the target sound was counterbalanced across all participants). During the presentation of the sequence, the participants were asked seven to nine times pseudorandomly during the sequence (i.e., with at least one trial between each question) to retrospectively report manually if the sound they heard was the "target" sound or not. The sequence stopped at each question and restarted as soon as the participant responded. Because the active listening condition was more demanding than the passive listening conditions, only 6 blocks of 1,000 stimuli were presented. The total numbers of presentation for each tone type were 1,766 (H), 1,837 (M), 1,785 (L), and 606 (O) in the active condition.

Critically, the onset of the next tone could be predicted in the sequence because the ISI was kept constant at 450 msec. The direction of the change in pitch of the next tones was also largely predictable in the sequence because, after a low-pitch tone, only a higher pitch tone (M or H) could be presented; after a high-pitch tone, only a lower pitch tone (M or L) could be presented; and after a medium-pitch tone, only a higher or lower

pitch tone could be presented (H or L). Although the precise pitch of the omission in the sequence could not be predicted, we could extract three types of omission, which differed in their expected pitch depending on the preceding pitch (see also the Discussion section).

Stimulus delivery and response recording were controlled by EPrime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

## EEG Analyses

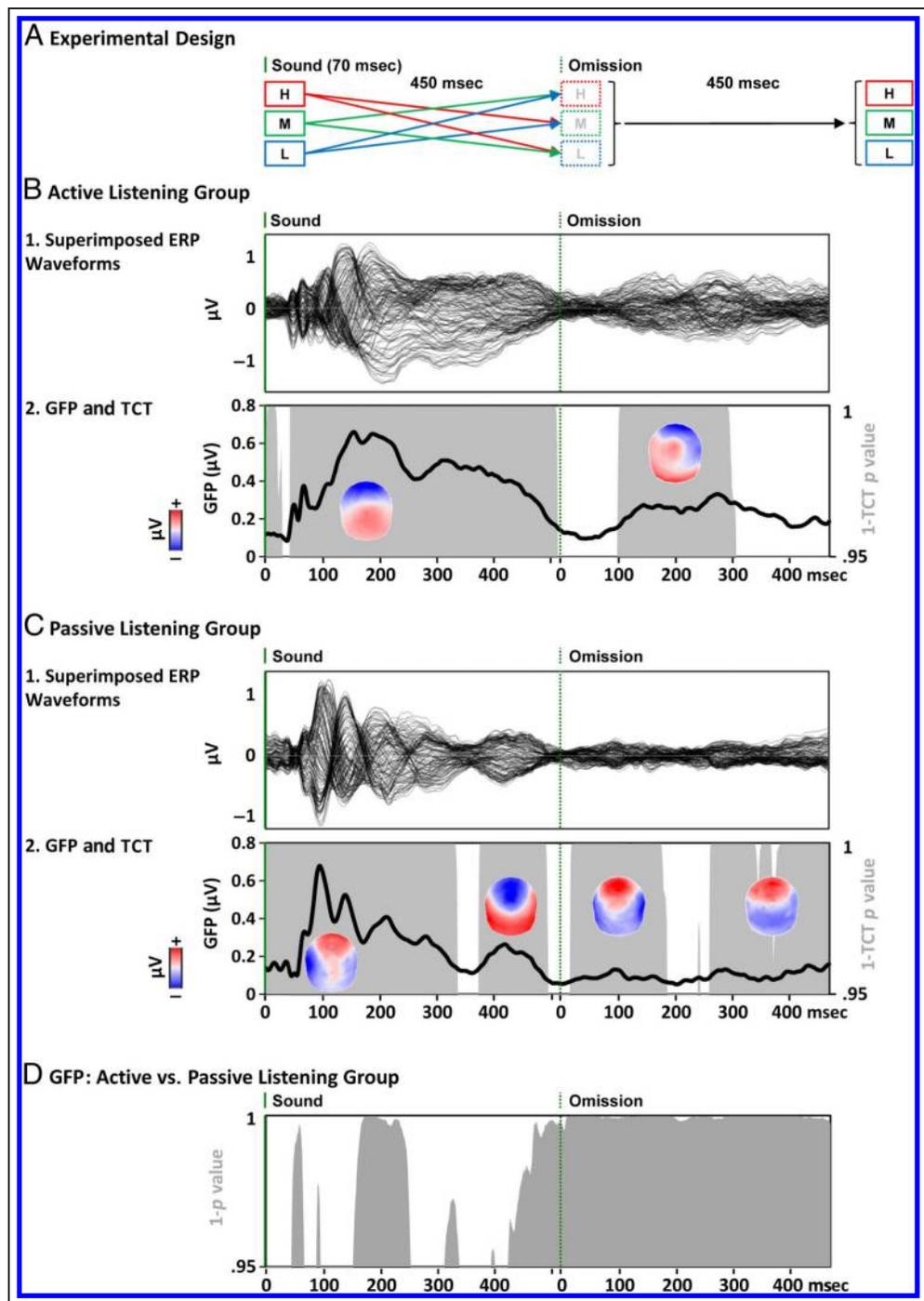
### *EEG Acquisition and Preprocessing*

Continuous EEG was acquired during the passive and active listening phases at 1024 Hz through a 128-channel Biosemi ActiveTwo system referenced to the common mode sense/driven right leg ground (which functions as a feedback loop driving the average potential across the montage as close as possible to the amplifier zero).

We focused on a period including the omitted tone and the immediately preceding real tone. EEG epochs from 520 msec before (corresponding to the onset of the sound preceding an omission, i.e., 450-msec ISI + 70-msec sound) to 470 msec after the onset of the omissions were extracted and were then averaged for each participant. Trials with blinks, eye movements, or transient noise were rejected using a semiautomated  $\pm 80\text{-}\mu\text{V}$  criterion and visual inspection. On average, for each participant, 11.7% of the trials were rejected for the high-pitch, 11.3% for the medium-pitch, and 13.8% for the low-pitch conditions. These values did not differ statistically ( $F(2, 23) = 0.742, p = .49$ ).

Four different ERPs were generated per participant. The first ERP included all omissions together, independent of the pitch of the tone preceding the omission (O; H-O + M-O + L-O), and was computed for the topographic consistency (TC) and global field power (GFP) analyses. These analyses aimed at characterizing the temporal dynamics of brain responses to omissions irrespective of their expected pitch. Including all pitch conditions in the ERP thus enabled improving the signal-to-noise ratio. For the topographic analyses, three other ERPs were generated, each depending on the tone preceding the omission (H-O, M-O, and L-O). Because the direction of the change in pitch of the forthcoming tone was predictable based on the preceding tone (at least for the after the H and L tones), these ERPs each included an omission with a different expected pitch. Before group averaging, artifactual electrodes from each participant (i.e., electrodes that were not considered during the trial rejections procedure because they showed transient noise higher than the  $\pm 80\text{-}\mu\text{V}$  criterion in more than five consecutive trials) were interpolated (an average of  $14.2 \pm 1.5$  and  $13.2 \pm 2.6$  electrodes [mean  $\pm$  SD] were interpolated in the active and passive groups, respectively; Perrin, Pernier, Bertrand, Giard, & Echallier, 1987), and then all electrodes were rereferenced against the average reference.

**Figure 1.** Experimental design and Topographic Consistency Tests (TCTs). (A) Experimental design. Passively or actively presented sound sequences were composed of high-frequency (H, 1100 Hz; in red), medium-frequency (M, 900 Hz; in green), or low-frequency (L, 700 Hz; in blue) 70-msec tones with a 450-msec ISI. Ten percent of the tones were pseudorandomly omitted during the sequences. When they were not followed by omissions, H tones were always followed by a lower pitch tone (M or L), M tones were always followed by a lower or higher pitch tone (L or H), and L tones were always followed by a higher pitch tone (M or H), so that the relative pitch (higher or lower) of the forthcoming tone was predictable. (B and C, top) Group-averaged ( $n = 13$ ) ERP waveforms from the 128 electrodes across participants. The ERPs collapsing all pitch conditions together are displayed in microvolts as a function of time. The ERP epoch of interest starts from the onset of the sound preceding the omission (plain green line) to 470 msec after the onset of the omission (dotted green line). (B and C, bottom) Global Field Power (GFP) and TCTs. The GFP is displayed in microvolts as a function of time (black line) and indexes the strength of the electric field at the scalp. The 1-TCT  $p$  value is plotted as a function of time and indicates the presence of an ERP signal significantly different from noise (gray areas,  $p < .05$ ) over almost all the periods after the onset of the real sound. There was also a significant consistency 100–310 (active listening group, B), 20–190, and 260–470 (passive listening group, C) msec after the onset of the omission. The topography of the ERP is represented (nasion upward; blue and red for negative and positive voltages, respectively) averaged over the period after the omission showing a significant TCT and after the corresponding period in the real sound. (D) GFP comparison between the active versus passive listening group. The 1- $p$  values are plotted as a function of time. The GFP of the ERP was higher (gray areas,  $p < .05$ ) in the active than passive listening condition 150–250 msec after the onset of the real sound and during the whole period after the omission.



## ERP Analyses

### Topographic Consistency Tests (TCTs)

As a first step, we investigated the spatio-temporal dynamics of the brain responses to omissions irrespective of the preceding tone by analyzing the topographic consistency of the ERP including all pitch conditions (passive

and active conditions separately) and the GFP between this ERP in the passive versus active listening condition. TCTs were carried out using the RAGU software package (Koenig, Kottlow, Stein, & Melie-Garcia, 2011), following the method developed by Koenig et al. (2011) and Koenig and Melie-Garcia (2010). The TCT aims at identifying the presence of a signal (i.e., an ERP) that is significantly

different from noise in the EEG data. The rationale behind the TCT is that, if there is a brain response functionally related to the omissions (i.e., if the brain responses at the moment of the omissions are not only noise), then this brain response should be similar across participants at a given latency after stimulus or omission onset. Because similar configurations of intracranial generators imply similar topographies, topographies are tested for consistency across participants to determine the presence of a response evoked by the omissions. The test for between-participant consistency is based on measures of the GFP. The GFP is calculated as the spatial standard deviation of the electric potentials, that is, as the square root of the sum of all squared potentials divided by the number of electrodes (Murray, Brunet, & Michel, 2008; Lehmann & Skrandies, 1980). The TCT analysis entails the comparison between the GFP of the group-averaged ERP at a given time frame (TF) versus the distribution of GFP values obtained after randomly reshuffling the voltage values measured at each electrode for each participant. It is worth noting that the GFP for each participant will not be affected by reshuffling whereas the GFP of the average will be sensitive to that procedure. The principle of this analysis is that the higher the consistency of the topographies across participants at a given latency, the higher GFP of the group-averaged ERP. Thus, the comparison between the GFP of the group-averaged ERP and the average of the GFP of the individual ERPs can be used as the effect size of the topographic consistency. On this basis, the topographic consistency is estimated as follows. For each participant and condition (i.e., passive and active listening), the electrodes are shuffled 5,000 times to generate a data set corresponding to a situation where the individual ERPs are only noise (i.e., the topographic information is destroyed while preserving the GFP). Then, the probability that the measured ERPs are only noise equals the percentage of the 5,000 randomizations in which the GFP in the group mean of the actual ERPs is higher than the GFP of the group mean of the shuffled data ( $p$  threshold was set at .05; see Tzovara, Murray, Michel, & De Lucia, 2012, for review).

### GFP Analyses

Modulations in the strength of the electric field at the scalp at each time point were quantified using the GFP (Koenig & Melie-Garcia, 2010; Murray et al., 2008; Lehmann & Skrandies, 1980). Differences in GFP between the active and passive listening conditions were assessed based on the same randomization procedure described for the TCT: GFP at each time point was compared with an empirical distribution of the GFPs derived from a bootstrapping procedure (5,000 permutations per data point) based on randomly reassigning each participant's data to one of the two listening conditions (Koenig et al., 2011; Koenig & Melie-Garcia, 2010). Because the ERP GFP is orthogonal to the ERP topography, GFP analyses enable as-

sessing modulations in response strength independently of the configuration of the underlying brain networks.

### Topographic Analyses

We investigated whether and when the expected pitch of an omitted tone had an effect on the electrophysiological response to the omission and whether this effect was different between the active and passive listening conditions. To do so, we compared the ERP topography with the omissions between the three pitch conditions and for the active and passive listening conditions using a  $3 \times 2$  mixed topographic ANOVA with pitch (three levels: H, M, and L) as within-participant factor and listening condition (active and passive) as between-participant factor. As for the TCT and the GFP analyses, randomization statistics were used for the topographic analysis at each time point over the whole EEG epoch.

The global ERP topographic analyses used in this study have several advantages over classical waveform analyses: They are reference independent (for details, see Tzovara, Murray, Michel, et al., 2012; Murray, Camen, Gonzalez Andino, Bovet, & Clarke, 2006; Murray et al., 2005; Michel et al., 2004; Lehmann, Ozaki, & Pal, 1987) and data driven (no a priori selection of electrodes or periods of interest).

Modulations in the topography of the electric field at the scalp at each time point were expressed as a dissimilarity index (global map dissimilarity [GMD]). For two experimental conditions, that is, only two levels of a factor, the GMD is the root mean square of the "difference" scalp map, that is, the electrode-wise difference in voltage potential. Importantly, before calculating the difference, each voltage was normalized by the GFP across all electrodes for each condition (for a review, see Murray et al., 2008).

The utilization of GFP-normalized topographic maps enables examination of topographic differences between maps independently of pure amplitude modulations. Because topographic changes necessarily follow from differences in the configuration of the brain's underlying active generators (Srebro, 1996; Lehmann et al., 1987), the analysis of GMD determines if and when different configurations of brain networks are engaged across experimental conditions. As with the TCT and GFP, to determine whether an observed GMD arose by chance, we used a randomization-based permutation test. When the experimental design includes only two conditions, the two topographic maps obtained for each participant at a given time point are randomly assigned to one of the two levels, and the grand-mean GMD for this permutation is obtained by averaging across participants. Once a sufficiently large number of permutations (maximum =  $2^N$ , where  $N$  is the number of participants) have been carried out to obtain an empirical distribution, the probability ( $p$ ) that the grand-mean GMD arose by chance can be expressed as the proportion of permutations in the reference distribution with a GMD greater than or equal to the

actual GMD. In the current experiment, we applied a  $3 \times 2$  Pitch (H, M, and L)  $\times$  Listening condition (active and passive) design for the topographic analyses, and so a generalized version of the above procedure was used (Koenig & Melie-Garcia, 2010; Wirth et al., 2008). First, the grand-mean scalp map was subtracted from every level of each participant to produce “residual” maps. Next, for each level, the residual maps were averaged across participants, and the GFP of this grand-mean residual map was calculated. A large GFP of the residual map indicates that the original map for this level differs greatly from the grand-mean map. Finally, the residual-map GFPs were summed across levels to provide a “generalized” GMD. Note that, in the case of only two levels, this is equivalent to the GMD described above. To obtain a reference distribution for generalized dissimilarity, the procedure was repeated 5,000 times with scalp maps randomly assigned to one of the levels in a within-participant manner. Calculations were performed with RAGU (Koenig et al., 2011), and  $p$  values for each time point are reported.

For the TCT, GFP, and GMD analyses, correction was made for temporal autocorrelation by considering only significant differential effects lasting for at least 11 contiguous data points (Guthrie & Buchwald, 1991).

Theoretically, significant topographic and GFP results should manifest only over periods of topographic consistency (i.e., when there is an ERP signal). These analyses should thus be carried out only over periods of significant TCT. However, with the aim of testing this assumption and of providing as much information as possible, we opted to conduct and show the results of these two analyses for each TF over the whole epoch.

### Electrical Source Estimations

A distributed linear inverse solution and the local autoregressive average (LAURA) regularization approach were used to estimate intracranial sources of the scalp-recorded data (Grave-de Peralta, Gonzalez-Andino, & Gomez-Gonzalez, 2004; Grave de Peralta Menendez, Murray, Michel, Martuzzi, & Gonzalez Andino, 2004). Intracranial sources (current densities at each solution point) were estimated and statistically processed over the periods showing significant pitch, listening condition, or Pitch  $\times$  Listening condition topographic modulation. ERPs for each participant and condition were first averaged over the period of interest. Then, intracranial sources were estimated for the resulting one time-sample ERP for each participant and condition. These processing steps were conducted using Cartool software (Brunet, Murray, & Michel, 2011). The current densities at each solution point were then statistically compared at each solution point between the experimental conditions using the same  $3 \times 2$  Pitch (H, M, and L)  $\times$  Listening condition (active and passive) ANOVA as for the ERP analyses. The solution space included 5,299 nodes equally distributed within the gray matter of the averaged brain of the Montreal Neurological

Institute (courtesy of Grave de Peralta Menendez and Gonzalez Andino, University Hospital of Geneva, Geneva, Switzerland). To control for multiple comparisons, only significant clusters with a minimal size of 14 consecutive points ( $k_E$ ) were retained. This spatial criterion was determined with the AlphaSim program (<http://afni.nimh.nih.gov/afni>); there was a false positive probability of  $<.005$  for observing a cluster of 14 nodes (see also Knebel & Murray, 2012; De Lucia, Clarke, & Murray, 2010). Statistical analyses in the brain space were performed using the STEN toolbox developed by Jean-François Knebel.

### Single-trial Classification

We used a single-trial classification technique to examine whether the ERP responses to omitted sounds share common features with the ERP responses to real sounds in both groups of participants. Specifically, we trained a classifier to discriminate single trials in response to higher (i.e., M and H tones following an L) versus lower (L and M tones following an H) real sound presentations and used this classifier to examine whether it could also discriminate EEG responses to higher (expected H or M, i.e., omissions following an L tone) versus lower (expected L or M = omissions following an H tone) omissions.

The classification technique implemented here consists of modeling the distribution of single-trial voltage topographies based on a mixture of Gaussians model (GMM). The GMM computation was based on the instantaneous measurements from all electrodes together, modeled as time points in an  $n$ -dimensional space, where  $n$  = total number of electrodes (here, 64). To ensure that any effect was a result of topographic but not amplitude modulation, before training the classifier, we normalized values on all electrodes by the instantaneous GFP.

The details of this method have been presented elsewhere (Tzovara, Murray, Bourdaud, et al., 2012; Tzovara, Murray, Michel, et al., 2012). This approach has been successfully used in the past to decode EEG responses to auditory stimuli (Cossy, Tzovara, Simonin, Rossetti, & De Lucia, 2014; Tzovara et al., 2013; De Lucia, Tzovara, Bernasconi, Spierer, & Murray, 2012; Bernasconi et al., 2011). Here, we applied this algorithm in the groups of passive and active participants separately. For each of the conditions, we extracted two nonoverlapping data sets, a cross-validation (CV) data set and a validation (V) data set. The CV data set consisted of 30 trials per participant, extracted randomly (390 trials in total) and corresponding to H and L real sound presentations. These trials were used for computing the GMM models and selecting the optimal parameters for classification in a 10-fold CV procedure.

The V data set consisted of 30 trials per participant, corresponding to 15 H and 15 L real sound responses (390 trials in total; validation real [VR] data set), and it was used for validating the computed models and examining their

accuracy for discriminating responses to real sounds. The choice of these proportions was based on our previous work (De Lucia et al., 2012; Bernasconi et al., 2011). Finally, we extracted 30 trials per participant (390 trials in total; validation omission data set) in response to 15 H and 15 L omissions to examine whether the classifiers trained on real sounds could also accurately classify responses to omissions.

In all of the abovementioned cases, we trained/tested and validated the classifier with instantaneous single-trial voltage topographies extracted from  $-100$ - up to  $500$ -msec poststimulus onset for real sounds or omissions. All the single trials were extracted randomly from all the blocks to cover the entire experiment duration.

The GMM models' computation was based on an expectation-maximization algorithm (Dempster, Laird, & Rubin, 1977) and was carried out for each experimental condition separately (i.e., responses to H vs. L real sounds). After estimating the GMM models, we assigned posterior probabilities on the single-trial voltage topographies, which represent the probability for every trial and time point to be represented by each of the Gaussians in the models (Tzovara, Murray, Michel, et al., 2012).

The GMM model computation was carried out on one part of the CV data set (training data set,  $\sim 90\%$  of the CV trials). The extracted features (i.e., GMM models and discriminative periods) were then used for classifying the remaining  $10\%$  of the CV trials (test data set) by computing posterior probabilities for each of the GMM models. To increase the signal-to-noise ratio of the single-trial EEG responses, we averaged four single trials belonging to each experimental condition and then classified this average. The classification performance was quantified as the area under the receiver operator characteristic curve (AUC; Green & Swets, 1966). The model computation and classification was repeated in a 10-fold CV, in a way such that the test data sets never overlapped.

As it is not possible to estimate in advance the total number of Gaussians in the GMM models, we trained a series of models for each condition, while varying the number of Gaussians from 3 to 10. Finally, we selected the pair of GMMs that provided the maximum AUC value on average on the 10 test data sets. Because these test data sets were used for model selection, a more realistic value of the decoding performance was obtained by classifying the validation trials using the selected models, which had never been used for training/testing the models. Here, we only report results corresponding to these validation trials.

In addition, we examined the generalization of the computed models to another group of participants: We used the models that were trained for the real sound presentation in the passive group to classify validation trials corresponding to omissions of the active group and vice versa. The goal of this cross-group analysis was to examine whether the computed models generalize not only to omissions from the same group of participants but also to omissions in a different group.

In all of the abovementioned cases, the significance of the classification results was assessed by comparing the classification performance on the validation data sets with chance level. Chance level was quantified by randomly permuting the true labels of the CV trials and re-computing the GMM models 500 times. These random models were then used for classifying the validation trials. The true decoding performance of the validation data sets, obtained with the original models, was compared with the distribution of the decoding values based on the random models (Wilcoxon signed-rank test,  $p < .001$ ). It is worth to note that comparing the true validation results with the chance level is equivalent to comparing it with a "nonpredictive" condition where the sound sequence information is lost.

## RESULTS

### ERP Analyses

#### TCTs

Figure 1B (top) and C (top) display the time course of the mean ERP (i.e., including all pitch conditions together) at each scalp electrode from the onset of the sound preceding the omission to  $470$  msec after the onset of the omission.

The results of the TCT are displayed in the Figure 1 (B and C, bottom). In both the active and passive listening conditions, the response to the real sound elicited periods of sustained significant topographic consistency, confirming that the TCT was sensitive to the ERP elicited by processing of real sound.

There were ERP signals significantly different ( $p < .05$ ;  $>11$  TFs) from noise also during the period after the onset of the omissions. In the active condition, there was a period of significant topographic consistency  $100$ – $310$  msec after the expected onset of the omission. In the passive condition, there were two periods of topographic consistency  $20$ – $190$  and  $260$ – $470$  msec after the expected onset of the omission.

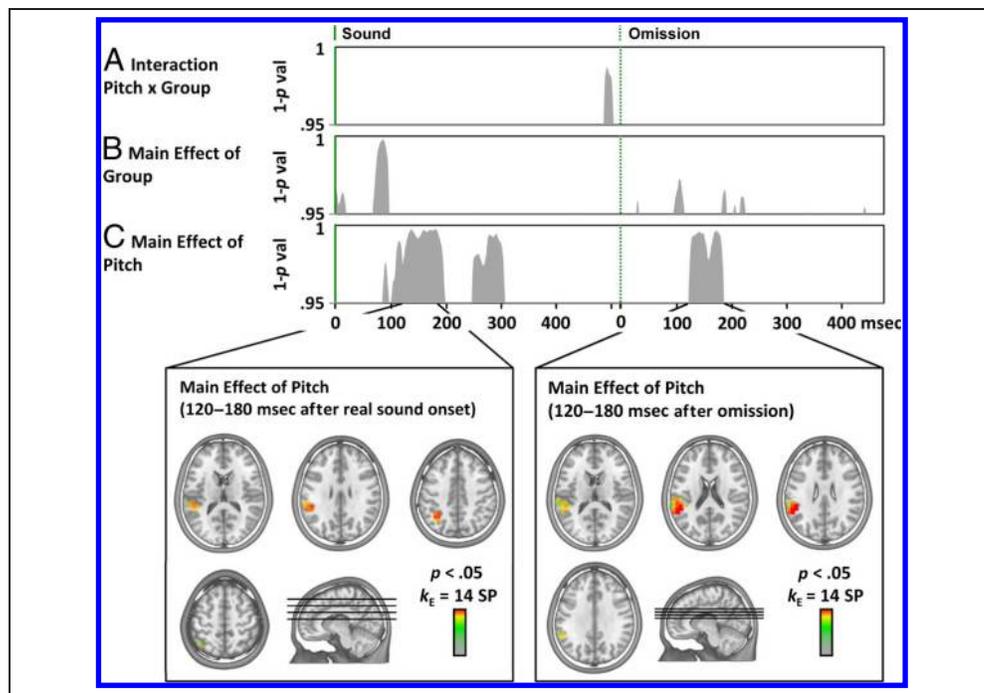
### GFP Analyses

The GFP of the ERP to the omissions was higher ( $p < .05$ ;  $>11$  TFs) in the active than passive listening condition  $150$ – $250$  msec after the onset of the real sound and, critically, over a large period of the omissions, indicating stronger global response strength to the omissions in the active than passive listening condition (Figure 1D).

### Topographic Analyses

The time-frame-wise  $3 \times 2$  Pitch (H, M, and L)  $\times$  Listening condition (active and passive) mixed topographic ANOVA revealed a main effect of factor Pitch  $85$ – $200$  and  $250$ – $310$  msec after the onset of the real sound. There was also a main effect of Pitch  $120$ – $180$  msec after the

**Figure 2.** Topographic ANOVA and source estimation (Pitch [H, M, and L] × Group [active and passive]) analyses. The 1-*p* value of the ERP topographic ANOVA is plotted as a function of time for the two main effects and the interaction term. There was a significant topographic main effect of Pitch 120–180 msec after the onset of the omission. A corresponding pattern was found in response to the real sound. Source estimation over the period of topographic modulation after the onset of the omission revealed that the topographic main effect of Pitch stemmed from left posterior temporal areas. The same analysis conducted over the corresponding period but after the onset of the real sound revealed a comparable pattern of result. Axial slices of the source estimation are centered on the voxel showing the most significant *p* value.



expected onset of the omission. There was a main effect of Group 70–100 msec after the onset of the real sound but not for the omissions ( $p < .05$ ; >11 TFs). There was no Pitch × Group interaction (Figure 2).

### Electrical Source Estimations

Electrical source estimations were submitted to a  $3 \times 2$  Pitch (H, M, and L) × Listening condition (active and passive) mixed ANOVA over periods showing a significant topographic modulation during omissions, that is, for the main effect of pitch 120–180 msec after the expected onset of the omission. The results revealed a significant main effect of Pitch within left posterior temporal regions ( $p < .05$ ;  $k_E = 14$  solution points). The same analysis was conducted over the same period but after the onset of the real sound. The pattern of result was very similar

and showed a main effect of Pitch within posterior temporal regions, with an additional involvement of the bilateral temporal poles (Figure 2).

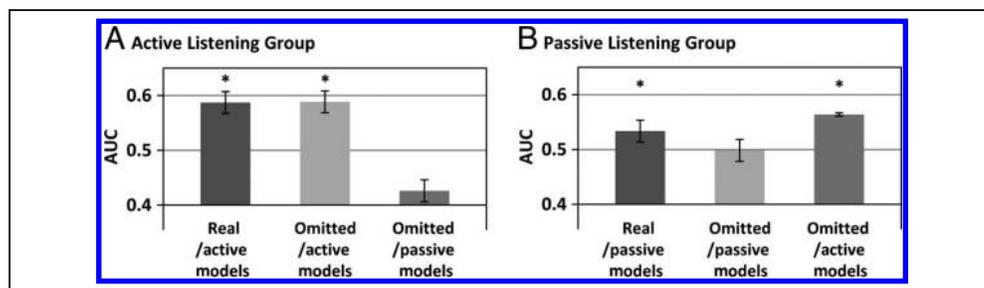
### Single-trial EEG Analyses

#### Classification of Responses to Real Sounds

In the active group of participants, classification performance of responses to H versus L sounds was  $0.67 \pm 0.07$  in the CV data set and  $0.59 \pm 0.01$  in a separate set of VR trials (AUC value ± standard error; Figure 3A). This result was significantly above chance level, based on a Wilcoxon signed-rank test ( $|z| = 18.38$ ,  $p < .001$ ). Chance level was equal to  $0.51 \pm 0.002$ .

In the passive group, classification performance of real sounds was  $0.73 \pm 0.08$  in the CV data set and  $0.53 \pm$

**Figure 3.** Decoding performance for the active and passive listening groups of participants obtained on the validation data set for each condition (mean and SEM shown). The decoding algorithm was trained using the models of the real presentations of H versus L pitch tones and could accurately discriminate EEG



responses to real sound presentations for both active and passive groups of participants (A and B, first bar). However, it could only discriminate omissions for the active group (A and B, second bar). The models of the active group were also able to generalize and discriminate responses of the passive group of participants, but not the other way around (A and B, third bar). Asterisks mark above-chance-level decoding results. AUC = area under the curve.

0.02 in the VR data set (Figure 3B). These results were also significantly above chance levels, although with a weaker  $z$  value than for the active group (Wilcoxon test:  $|z| = 9.07, p < .001$ ). Chance level was equal to  $0.51 \pm 0.003$ .

The above-chance-level classification results on the real sounds suggest that the computed models were able to reliably extract informative features of the single-trial EEG responses to real presentations of H versus L pitch sounds. In summary, these results suggest that the models for the active group of participants were more robust compared with the passive one, as indicated by a higher AUC value in the VR data set and higher level of significance. We further tested whether the same models, computed for real sounds, can be generalized for classifying responses to omissions.

### *Classification of Responses to Omissions*

In this context of omitted responses, we used the same models as for the real sound presentations. These models were obtained based on the CV data set for which results have been reported above. In the active group, classification performance for omissions in the validation data set was  $0.59 \pm 0.01$  (Figure 3A). This result was significantly above chance levels (Wilcoxon signed-rank test:  $|z| = 18.45, p < .001$ ), suggesting that responses to omissions shared common features with responses to real sounds and that these features are specific to the expected pitch. Chance level was  $0.50 \pm 0.003$ .

In the passive group, however, classification of omissions in the validation data set was around chance levels, with an AUC value of  $0.50 \pm 0.02$  (Figure 3B), and we therefore did not compute the random permutations. These classification results of the omissions for both groups of participants were based on the same models as the results for the real sound presentations.

In addition, we examined whether the computed models, based on the real sounds of one group of participants, could also be generalized to the other group. We used the models of the passive group to classify omissions from the active group. The classification performance in this case was very poor and equal to  $0.43 \pm 0.02$  (Figure 3A). By contrast, when using the models of the participants in the active group, we were able to accurately classify the omissions of the passive group: The classification performance in this case was  $0.56 \pm 0.003$  (Figure 3B), and chance level was  $0.52 \pm 0.003$ . These results were significantly above chance level (Wilcoxon signed-rank test:  $|z| = 14.24, p < .001$ ).

We also examined whether the residual state from the preceding sounds was driving the classification. If that was the case, then the classifier should work when considering only the preevent period as it does when considering the whole trial, including the post-event interval, because residual state before real sound is the same as before omissions. On the basis of this rationale, we tested

the classification based on the models of the active conditions but taking into account only 25 msec preceding the onset of the active and passive omissions for the testing of the classifier. This analysis resulted in classification performance at chance level (passive: 0.48 AUC, active: 0.48 AUC), indicating that residual state from the preceding sound did not drive above-chance classification and thus cannot account for our results.

## **DISCUSSION**

To determine whether and how experience-based auditory internal models generate brain activity during predicted but omitted stimuli, we analyzed brain responses to omitted sounds in passively or actively presented rule-based sequences of tones with varying pitches.

We found an above-chance-level decoding of the electrophysiological responses to the omissions in both the passive and active listening conditions when decoding was based on the responses to real sounds in the active condition. Although the decoding of the real sounds in the passive condition was also above chance level for models based on real sounds in this condition, models based on these brain responses could not decode responses to omissions in either the passive or active conditions.

Topographic consistency analyses after the onset of the omitted tones revealed the presence of an electrophysiological response even when an expected tone was not presented. ERPs to omissions manifested from 50 msec after the expected onset of the omission in the passive and from 100 msec in the active listening condition. Overall, the electrophysiological response to the omissions was stronger in the active than passive listening condition, which corresponds to the effects of attention usually observed on the processing of real sounds (e.g., Coch, Sanders, & Neville, 2005).

The Pitch  $\times$  Listening topographic analyses revealed a main effect of Pitch around 100 msec after the onset of the omission, indicating that different brain networks were engaged in response to the omission depending on its expected pitch. Source estimations revealed that this topographic modulation stemmed from the left posterior temporal cortices.

Although previous literature reports “similar” (SanMiguel, Widmann, et al., 2013; Bendixen et al., 2009) or “correlated” responses for omissions and actual sounds (Janata, 2001), there was so far no direct positive statistical evidence that neural responses to predicted but omitted sounds match responses to the actual sounds that have been omitted. This lack of evidence pertains to the fact that the classical null hypothesis testing approach can only state a failure to reject the null hypothesis (e.g., Rouder et al., 2009). Null results can indeed always be argued to follow from data insensitivity because of a lack of statistical power, weaknesses in the experimental design, or other limitations. To circumvent this problem, we applied

a classification analysis on the single-trial ERPs to the actual sounds and to the omissions. We found an above-chance decoding of the electrophysiological responses to the omission in passive and active listening conditions, in both cases, by training responses to the real sounds of the active condition. This pattern of results provides the first direct statistical evidence that the pitch-specific topography—and thus, the brain network—to the actual sounds matches those of the omissions. The accuracy of these decoding results is particularly relevant given that they are based on single-trial responses across all the participants in each of the active and passive conditions, thus showing (i) a high degree of generalization of the response pattern and (ii) that a similar neural correlate of the predictive phenomenon is indeed present across individuals.

We found that the prediction process took place even during the passive listening condition, corroborating that the building up of internal models is an automatic stimulus-driven mechanism (Bendixen et al., 2009). However, whereas responses to the real sounds in the passive condition were decodable, this was not the case for responses to the omissions. This finding suggests that attention might have influenced the sound features encoded in internal models. However, given that the strength of the response was weaker in the passive condition (as indicated by the GFP analysis), one explanation could be that the classifier was penalized in a context of a low signal-to-noise ratio although the pitch features were still encoded.

The fact that brain responses to omissions contained pitch information matching those in the responses to actual sounds confirms that brain activity from internal models can be considered as contributing to auditory filling-in phenomena such as phoneme restoration (Warren, 1970) or continuity illusions (Miller & Licklider, 1950). Indeed, it has long been known that an interrupted signal, such as a pure tone or speech, can be perceived as continuous if the interruptions are “filled in” by a masker, such as broadband noise (Houtgast, 1972; Elfner & Caskey, 1965). Importantly, this continuity illusion occurs only if the masker contains sufficient energy at the signal frequency. The effect of frequency in the current experiment (main effect of pitch) suggests a potential neural mechanism by which predictive mechanisms may contribute to the continuity phenomenon or, more broadly, to the auditory imagery evoked by silent gaps in musical pieces (Margulis, 2007): They would produce activity mimicking responses to actual sounds as in this study. This would imply that auditory cortical activity is necessary, but not sufficient, for the conscious perception of a sound as “real” and has implications for our understanding of auditory imagery and its related disorders such as tinnitus or verbal and musical hallucinations. However, phoneme restoration and gap filling-in take place over a much shorter time scale than the processes involved in our studies. In addition, our design incorporates rhythmic aspects that are not necessarily involved in the previous literature on filling-in phenomena.

The time-frame-wise TCT revealed the precise temporal dynamics of brain responses to an expected but omitted tone. Periods of topographic consistency, indicating when there was an ERP significantly different from noise in response to an omission, started earlier in the passive listening condition (ca. 50 msec) than in the active listening condition (ca. 100 msec). These results corroborate and extend previous ERP omission studies, which identified two main ERP components signaling predictive activity in response to omissions in sound sequences (for a review, Bendixen et al., 2012). As with our results, a P50 component to omissions was observed at 50 msec by Bendixen and colleagues (2009) and Tervaniemi, Saarinen, Paavilainen, Danilova, and Näätänen (1994) in passive listening conditions, and an N1 component was observed at 100 msec by Janata (2001) in active listening conditions.

We extend these findings by revealing that, in active listening conditions, there was no brain response to the omission before 100 msec and that an ERP is present until 300 msec, thus covering the classical omission-related MMN period in addition to the N1 (Yabe et al., 1998; Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997). In contrast, the latency of initial response to an omission was shifted 50 msec earlier in the passive listening condition and covered a period corresponding to the middle latency components (Yvert, Crouzeix, Bertrand, Seither-Preisler, & Pantev, 2001). Our results thus suggest that attention actually delays the latency of brain responses to omissions. This shift may reflect a focus on specific task-relevant features of the sound (i.e., a specific pitch in our active listening condition). In contrast, early detection of new, or of a change in forthcoming, auditory information might be favored in passive listening conditions to call rapidly attention to unattended auditory events, independent of their specific features. Of note, this pattern of results seems inconsistent with studies on the effects of attention on responses to real sounds. Folyi, Feher, and Horvath (2012) showed that attention speeds up early auditory processing, as demonstrated by shorter N1 component latency to sounds in a sound detection versus an “ignore” condition (Folyi et al., 2012).

The comparison between the GFP to the omissions in the passive versus active listening conditions revealed that active listening increased response strength to the omission. This result suggests that, independent of the configuration of the brain network engaged in the response to omissions, attention modulates their response gain. In addition, the fact that attention modulates the response strength to omissions in the same way as for real sounds further supports that internal models generate brain activity mimicking those elicited by real stimulations. We would note that, because the active and passive conditions refer to different participants, our effects most likely reflect differences in sustained attention and could be merely because of differences in task difficulty. Supporting this hypothesis, transient shifts in attention rather

manifest as auditory N1 amplitude gain, whereas our results show long-lasting GFP differences. Second, because twice as many stimuli were presented in the passive condition relative to the active condition, repetition suppression might well account for the decrease in power observed in the passive condition. Finally, we cannot rule out from our between-participant design that individual differences might also account for our results. Because target detection processes were present in the active condition but not in the passive condition and the target tones were included in the analyses of the active condition, target detection processes may also have participated in the GFP difference and confounded the effects of attention.

The time-frame-wise  $3 \times 2$  Pitch (H, M, and L)  $\times$  Listening condition (active and passive) mixed topographic ANOVA revealed a main effect of factor Pitch at 100 msec in response to omissions, driven by a modulation of left posterior temporal areas. As also corroborated by the evidence for a main effect of Pitch over the corresponding period and the same location in response to the real sounds in the present data, this main effect in response to omissions corresponds to pitch-sensitive latency and brain regions (e.g., Griffiths et al., 2010; Hyde, Peretz, & Zatorre, 2008; Schonwiesner & Zatorre, 2008; Seither-Preisler, Patterson, Krumbholz, Seither, & Lutkenhoner, 2006; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Griffiths, Uppenkamp, Johnsrude, Josephs, & Patterson, 2001; Lutkenhoner, Lammertmann, & Knecht, 2001; Griffiths, Buchel, Frackowiak, & Patterson, 1998; Zatorre, Evans, & Meyer, 1994; Zatorre, Evans, Meyer, & Gjedde, 1992). Importantly, there was no evidence for a main effect of Pitch over a 300-msec period before our effect, suggesting that it did not follow from mere differences in baseline because of the processing of the preceding real sound (Bendixen et al., 2012). Because frequency tuning curves are broad at the level of large populations of neurons, it may seem surprising that our electrical neuroimaging methods detected variations in brain responses to pitch differences of 200–300 Hz (Saenz & Langers, 2014). Our main effect of Pitch at 100 msec within posterior temporal cortices in responses to omitted tones thus unlikely reflects the very initial stages of frequency processing but rather secondary, associative processes related to perceiving tones of different frequencies. This hypothesis is further supported by our finding that this modulation was also found in response to the real sounds, which differed only at the level of their frequency.

Although negative results should be interpreted with caution, the absence of a Pitch  $\times$  Group interaction suggests that internal models impact on pitch-sensitive cortical activity similarly in active versus passive listening condition.

We would note that our study suffers several limitations that call for further investigation. First, because omissions also constitute a violation of a participant's expectations on the presentation of a sound during the se-

quence (different expected vs. actual sensory input), the EEG signal measured during the omission may also reflect an error response, which could likewise be modulated by frequency. The brain activity measured during the omission may thus reflect not only responses mimicking those to real sounds with corresponding features but also error signals. Although the current design cannot disentangle the relative contribution between these two factors, our pattern of results suggests that error response unlikely accounts for all our effects. The responses to the omission were indeed decodable from the response to the real sounds; because the response to the real sound does not contain error detection component, if the response to the omission were solely related to an error detection process, the decoding would have probably failed. In the same vein, the pitch modulation at 100 msec within posterior temporal cortices during the real sound was very similar (at both the spatial and temporal levels) to the one during the omission, suggesting that corresponding processes took place in the two conditions. As a third line of evidence speaking in favor of our interpretation, the latency of error detection processes such as the MMN generally takes place slightly later than our effects, at 150–200 msec (Fishman, 2014).

The fact that the direction of the change in pitch of the next tone, and not the pitch per se, was predictable in our sequence (e.g., an M or an L tone could follow an H tone) and that this direction was not predictable after an M tone constitute another limitation of this study. However, the levels of our factor pitch can still be interpreted in terms of differences in expected pitch. We chose to include the omission after an M tone in our ERP analyses to account for our full data set. This choice was actually conservative because, in the worst case, including the M tone would have increased noise in the data and thus the probability of Type II errors. As a control, we nonetheless also conducted the topographic ERP analyses without the M conditions and found the same pattern of results, that is, a main effect of pitch at 100 msec after the onset of the omission and no main effect of group or interaction (unpresented data).

The results of this study encourage further investigation of the capacity of the human brain for predicting incoming information in more challenging contexts, such as those in which the auditory sequences are based on rules organized over longer time scales than the single sound, based on specific auditory features or acoustic rhythms. These future studies will ultimately impact our understanding of the neural mechanism underlying the ability to process and perceive music and speech in noisy environments in a daily life context.

## Acknowledgments

This work was supported by a grant from the Service "Projets et Organisation Stratégiques" of the University Hospital of Lausanne to M.DL. (#29062-1144) and by a grant from the

Swiss National Science Foundation to L. S. (320030-143348). We thank Skander Mensi for his help in generating the sequences of stimuli and Michael Mouthon for his help in collecting and analyzing the data. Cartool software has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory, Geneva, Switzerland, and supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne. The STEN toolbox has been programmed by Jean-François Knebel, from the Laboratory for Investigative Neurophysiology (the LINE), Lausanne, Switzerland, and is supported by the Center for Biomedical Imaging (CIBM) of Geneva and Lausanne and by National Center of Competence in Research project “SYNAPSY—The Synaptic Bases of Mental Disease” (Project no. 51AU40\_125759).

Reprint requests should be sent to Dr. Lucas Spierer, Neurology Unit, Medicine Department, Faculty of Science, University of Fribourg, Ch. Du Musée 5, CH-1700, Fribourg, Switzerland, or via e-mail: Lucas.spierer@unifr.ch.

## REFERENCES

- Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. *Trends in Cognitive Sciences*, *10*, 93–94.
- Bendixen, A., SanMiguel, I., & Schroger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*, 120–131.
- Bendixen, A., Schroger, E., & Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. *Journal of Neuroscience*, *29*, 8447–8451.
- Bernasconi, F., De Lucia, M., Tzovara, A., Manuel, A. L., Murray, M. M., & Spierer, L. (2011). Noise in brain activity engenders perception and influences discrimination sensitivity. *Journal of Neuroscience*, *31*, 17971–17981.
- Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, *2011*, 813870.
- Coch, D., Sanders, L. D., & Neville, H. J. (2005). An event-related potential study of selective auditory attention in children and adults. *Journal of Cognitive Neuroscience*, *17*, 605–622.
- Cossy, N., Tzovara, A., Simonin, A., Rossetti, A. O., & De Lucia, M. (2014). Robust discrimination between EEG responses to categories of environmental sounds in early coma. *Frontiers in Psychology*, *5*, 155.
- De Lucia, M., Clarke, S., & Murray, M. M. (2010). A temporal hierarchy for conspecific vocalization discrimination in humans. *Journal of Neuroscience*, *30*, 11210–11221.
- De Lucia, M., Tzovara, A., Bernasconi, F., Spierer, L., & Murray, M. M. (2012). Auditory perceptual decision-making based on semantic categorization of environmental sounds. *NeuroImage*, *60*, 1704–1715.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, *39*, 1–38.
- Dubnov, S. (2008). Unified view of prediction and repetition structure in audio signals with application to interest point detection. *IEEE Transactions on Audio, Speech and Language Processing*, *16*, 327–337.
- Elfner, L., & Caskey, W. E. (1965). Continuity effects with alternately sounded noise and tone signals as a function of manner of presentation. *Journal of the Acoustical Society of America*, *38*, 543–547.
- Fishman, Y. I. (2014). The mechanisms and meaning of the mismatch negativity. *Brain Topography*, *27*, 500–526.
- Folyi, T., Feher, B., & Horvath, J. (2012). Stimulus-focused attention speeds up auditory processing. *International Journal of Psychophysiology*, *84*, 155–163.
- Grave-de Peralta, R., Gonzalez-Andino, S., & Gomez-Gonzalez, C. M. (2004). The biophysical foundations of the localisation of encephalogram generators in the brain. The application of a distribution-type model to the localisation of epileptic foci. *Revista de Neurologia*, *39*, 748–756.
- Grave de Peralta Menendez, R., Murray, M. M., Michel, C. M., Martuzzi, R., & Gonzalez Andino, S. L. (2004). Electrical neuroimaging based on biophysical constraints. *NeuroImage*, *21*, 527–539.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Griffiths, T. D., Buchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, *1*, 422–427.
- Griffiths, T. D., Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., et al. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, *20*, 1128–1132.
- Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience*, *4*, 633–637.
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*, 240–244.
- Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America*, *51*, 1885–1894.
- Hyde, K. L., Peretz, I., & Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*, *46*, 632–639.
- Janata, P. (2001). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, *13*, 169–193.
- Knebel, J. F., & Murray, M. M. (2012). Towards a resolution of conflicting models of illusory contour processing in humans. *NeuroImage*, *59*, 2808–2817.
- Koenig, T., Kottlow, M., Stein, M., & Melie-Garcia, L. (2011). Ragu: A free tool for the analysis of EEG and MEG event-related scalp field data using global randomization statistics. *Computational Intelligence and Neuroscience*, *2011*, 938925.
- Koenig, T., & Melie-Garcia, L. (2010). A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topography*, *23*, 233–242.
- Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, *434*, 158.
- Lehmann, D., Ozaki, H., & Pal, I. (1987). EEG alpha map series: Brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and Clinical Neurophysiology*, *67*, 271–288.
- Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609–621.
- Lutkenhoner, B., Lammertmann, C., & Knecht, S. (2001). Latency of auditory evoked field deflection N100m ruled by pitch or spectrum? *Audiology & Neuro-Otology*, *6*, 263–278.
- Margulis, E. H. (2007). Silences in music are musical not silent: An exploratory study of context effects on the experiences of

- musical pauses. *Music Perception: An Interdisciplinary Journal*, *24*, 485–506.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, *115*, 2195–2222.
- Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, *22*, 167–173.
- Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, *20*, 249–264.
- Murray, M. M., Camen, C., Gonzalez Andino, S. L., Bovet, P., & Clarke, S. (2006). Rapid brain discrimination of sounds of objects. *Journal of Neuroscience*, *26*, 1293–1302.
- Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., et al. (2005). Grabbing your ear: Rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex*, *15*, 963–974.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*, 2544–2590.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, *36*, 767–776.
- Perrin, F., Pernier, J., Bertrand, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and Clinical Neurophysiology*, *66*, 75–81.
- Raij, T., McEvoy, L., Makela, J. P., & Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain Research*, *745*, 134–143.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin and Review*, *16*, 225–237.
- Saenz, M., & Langers, D. R. (2014). Tonotopic mapping of human auditory cortex. *Hearing Research*, *307*, 42–52.
- Sanmiguel, I., Saupe, K., & Schroger, E. (2013). I know what is missing here: Electrophysiological prediction error signals elicited by omissions of predicted “what” but not “when”. *Frontiers in Human Neuroscience*, *7*, 407.
- SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schroger, E. (2013). Hearing silences: Human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, *33*, 8633–8639.
- Schonwiesner, M., & Zatorre, R. J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. *Experimental Brain Research*, *187*, 97–105.
- Schroger, E., Bendixen, A., Trujillo-Barreto, N. J., & Roeber, U. (2007). Processing of abstract rule violations in audition. *PLoS One*, *2*, e1131.
- Seither-Preisler, A., Patterson, R., Krumbholz, K., Seither, S., & Lutkenhoner, B. (2006). Evidence of pitch processing in the N100m component of the auditory evoked field. *Hearing Research*, *213*, 88–98.
- Shinozaki, N., Yabe, H., Sato, Y., Hiruma, T., Sutoh, T., Matsuoka, T., et al. (2003). Spectrotemporal window of integration of auditory information in the human brain. *Brain Research. Cognitive Brain Research*, *17*, 563–571.
- Srebro, R. (1996). An iterative approach to the solution of the inverse problem. *Electroencephalography and Clinical Neurophysiology*, *98*, 349–362.
- Stekelenburg, J. J., & Vroomen, J. (2015). Predictive coding of visual-auditory and motor-auditory events: An electrophysiological study. *Brain Research*. [Epub ahead of print].
- Sussman, E. S. (2007). The role of context in processing auditory events. *Journal of Psychophysiology*, *21*, 164–175.
- Tervaniemi, M., Saarinen, J., Paavilainen, P., Danilova, N., & Näätänen, R. (1994). Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biological Psychology*, *38*, 157–167.
- Tzovara, A., Murray, M. M., Bourdaud, N., Chavarriaga, R., Millan Jdel, R., & De Lucia, M. (2012). The timing of exploratory decision-making revealed by single-trial topographic EEG analyses. *Neuroimage*, *60*, 1959–1969.
- Tzovara, A., Murray, M. M., Michel, C. M., & De Lucia, M. (2012). A tutorial review of electrical neuroimaging from group-average to single-trial event-related potentials. *Developmental Neuropsychology*, *37*, 518–544.
- Tzovara, A., Rossetti, A. O., Spierer, L., Grivel, J., Murray, M. M., Oddo, M., et al. (2013). Progression of auditory discrimination based on neural decoding predicts awakening from coma. *Brain*, *136*, 81–89.
- Voisin, J., Bidet-Caulet, A., Bertrand, O., & Fonlupt, P. (2006). Listening in silence activates auditory areas: A functional magnetic resonance imaging study. *Journal of Neuroscience*, *26*, 273–278.
- 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, U.S.A.*, *108*, 20754–20759.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 392–393.
- Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, *742*, 239–252.
- Wirth, M., Horn, H., Koenig, T., Razafimandimby, A., Stein, M., Mueller, T., et al. (2008). The early context effect reflects activity in the temporo-prefrontal semantic system: Evidence from electrical neuroimaging of abstract and concrete word reading. *Neuroimage*, *42*, 423–436.
- Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport*, *8*, 1971–1974.
- Yabe, H., Tervaniemi, M., Sinkkonen, J., Huutilainen, M., Ilmoniemi, R. J., & Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. *Psychophysiology*, *35*, 615–619.
- Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., & Pantev, C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cerebral Cortex*, *11*, 411–423.
- Zanto, T. P., Snyder, J. S., & Large, E. W. (2006). Neural correlates of rhythmic expectancy. *Advances in Cognitive Psychology*, *2*, 221–231.
- Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908–1919.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846–849.

**This article has been cited by:**

1. Marc Recasens, Peter J. Uhlhaas. 2017. Test-retest reliability of the magnetic mismatch negativity response to sound duration and omission deviants. *NeuroImage* **157**, 184-195. [[Crossref](#)]
2. Elsa Juan, Nathalie Ata Nguenjo Nguissi, Athina Tzovara, Dragana Viceic, Marco Rusca, Mauro Oddo, Andrea O. Rossetti, Marzia De Lucia. 2016. Evidence of trace conditioning in comatose patients revealed by the reactivation of EEG responses to alerting sounds. *NeuroImage* **141**, 530-541. [[Crossref](#)]