



HAL
open science

Cortical responses to looming sources are explained away by the auditory periphery

Sarah Benghanem, Rudradeep Guha, Estelle Pruvost-Robieux, Julie Lévi-Strauss, Coralie Joucla, Alain Cariou, Martine Gavaret, Jean-Julien Aucouturier

► **To cite this version:**

Sarah Benghanem, Rudradeep Guha, Estelle Pruvost-Robieux, Julie Lévi-Strauss, Coralie Joucla, et al.. Cortical responses to looming sources are explained away by the auditory periphery. *Cortex*, 2024, 177, pp.321-329. 10.1016/j.cortex.2024.05.018 . hal-04788965

HAL Id: hal-04788965

<https://hal.science/hal-04788965v1>

Submitted on 18 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Research Report

Cortical responses to looming sources are explained away by the auditory periphery



Sarah Benghanem^{a,b,c,*}, Rudradeep Guha^d, Estelle Pruvost-Robieux^{a,c,e},
Julie Lévi-Strauss^e, Coralie Joucla^{d,f}, Alain Cariou^{b,c},
Martine Gavaret^{a,c,e} and Jean-Julien Aucouturier^d

^a INSERM UMR 1266, IPNP (Institut de Psychiatrie et Neurosciences de Paris), Paris, France

^b Medical ICU, Cochin Hospital, AP-HP, Paris, France

^c University Paris Cité, Medical School, Paris, France

^d Université de Franche-Comté, SUPMICROTECH, CNRS, institut FEMTO-ST, Besançon, France

^e Neurophysiology Department, GHU Paris Psychiatrie et Neurosciences, Sainte Anne Hospital, Paris, France

^f Science & Technologies of Music & Sound (STMS), IRCAM/CNRS/Sorbonne Université, Paris, France

ARTICLE INFO

Article history:

Received 6 November 2023

Reviewed 11 March 2024

Revised 14 May 2024

Accepted 22 May 2024

Action editor Veena Kumari

Published online 11 June 2024

Keywords:

Event-related potentials

Mismatch negativity

Oddball paradigm

Looming sounds

Cortical responses

ABSTRACT

A wealth of behavioral evidence indicate that sounds with increasing intensity (i.e. appear to be looming towards the listener) are processed with increased attentional and physiological resources compared to receding sounds. However, the neurophysiological mechanism responsible for such cognitive amplification remains elusive. Here, we show that the large differences seen between cortical responses to looming and *receding* sounds are in fact almost entirely explained away by nonlinear encoding at the level of the auditory periphery. We collected electroencephalography (EEG) data during an *oddball paradigm* to elicit mismatch negativity (MMN) and others Event Related Potentials (ERPs), in response to deviant stimuli with both dynamic (looming and receding) and constant level (flat) differences to the standard in the same participants. We then combined a computational model of the auditory periphery with generative EEG methods (temporal response functions, TRFs) to model the single-participant ERPs responses to flat deviants, and used them to predict the effect of the same mechanism on looming and receding stimuli. The flat model explained 45% variance of the looming response, and 33% of the receding response. This provide striking evidence that difference wave responses to looming and *receding* sounds result from the same cortical mechanism that generate responses to constant-level deviants: all such differences are the sole consequence of their particular physical morphology getting amplified and integrated by peripheral auditory mechanisms. Thus, not all effects seen cortically proceed from top-down modulations by high-level decision variables, but can rather be performed early and efficiently by feed-forward peripheral mechanisms that evolved precisely to sparing subsequent networks with the necessity to implement such mechanisms.

* Corresponding author. INSERM UMR 1266, IPNP (Institut de Psychiatrie et Neurosciences de Paris), Paris, France.

E-mail address: sarah.benghanem@aphp.fr (S. Benghanem).

<https://doi.org/10.1016/j.cortex.2024.05.018>

0010-9452/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The human auditory system has evolved to respond efficiently to fast and unpredictable changes in the acoustic environment that could be relevant for survival. One of the most salient examples of such prioritized auditory processing is the perceptual bias towards looming vs receding sound sources, which are typically simulated in the lab using simple increasing or decreasing changes of intensity sound level (Kolarik et al., 2016). The saliency of looming source produced by increasing intensity sound levels is a hallmark of human psychoacoustics: participants consistently overestimate the loudness (Neuhoff, 1998; Ponsot et al., 2015) and speed (Rosenblum et al., 1987; Schiff & Oldak, 1990) of looming compared to receding sounds. Physiologically, looming sounds also elicit stronger orienting response measured by skin conductance and heart rate changes (Bach et al., 2008; Bach et al., 2009; Tajadura-Jiménez et al., 2010), and facilitate the processing of associated visual stimuli (Leo et al., 2011; Romei et al., 2009). Finally, brain imaging studies have shown that looming and receding sounds activate brain areas related to spatial auditory processing, which include the right temporal plane and the right superior temporal sulcus (Alho et al., 2014; Seifritz et al., 2002), and that looming sounds activate a wider network of regions subserving auditory spatial perception and attention compared to receding sounds, including the right amygdala and left temporal areas (Bach et al., 2008; Seifritz et al., 2002). In sum, a wealth of behavioral and brain-imaging evidence converges to indicate that sound with increasing intensity function is an elementary warning cue, able to elicit adaptive responses by recruiting additional attentional and physiological resources.

Despite all this, event-related potential (ERP) evidence for the prioritized or amplified processing of looming vs receding sounds have remained remarkably contrasted. When comparing changing-level sounds with more frequent constant-level standards, some studies have documented earlier and higher mismatch negativity (MMN) for looming than for receding sounds, which is coherent with the general pattern of “cognitive amplification” of looming sounds (Shestopalova et al., 2018); but others have found that MMN amplitude increase with the magnitude of intensity change irrespective of direction (Rinne et al., 2006); and several studies have also reported no differences in MMN latencies or amplitudes for either looming or receding sounds (Altmann et al., 2013; Näätänen et al., 1993).

On closer inspection, these different results were obtained with experimental stimuli which, despite sharing a general pattern of increasing or decreasing amplitude, actually display a wide diversity of temporal characteristics (stepwise or gradual changes, linear or exponential profiles, duration and onset of level ramp). Psychophysical research has highlighted that loudness integration in changing-level sounds is not

identically distributed in time, and that perceptive weights are biased towards the beginning and end of the sound (Ponsot et al., 2013). If ERPs reflect such integration, then the latency and amplitude of the MMN responses may depend heavily on the morphology of the deviant. In addition, electrophysiological research shows that, even in mice, populations of neurons in the auditory cortices respond to rising or decreasing intensity-ramps asymmetrically, as the direct result of neuronal adaptation and non-linearities in their temporal integration (Deneux et al., 2016). For all these reasons, very large differences between looming vs receding MMN responses do not necessarily indicate, as often implied in the literature, a top-down cognitive amplification or prioritization of one type of sound over the other, but could be the result of the bottom-up integration of complex temporal profiles in stimuli and temporal non-linearities in their subsequent processing, i.e. of the same generic sensory mechanisms that would generate MMN to e.g. unremarkable constant-level deviants.

To clarify whether and how ERP responses to looming vs receding sounds coincide with behavioral and imaging evidence of their cognitive amplification, we need a way to model sensory contributions of time-varying stimuli to evoked potentials responses, and explore how much these generic mechanisms explain responses to specifically looming and receding sounds. To do this, we collected scalp EEG MMN data in response to deviant stimuli that presented potentially both dynamic (looming and receding) and constant level (flat) differences to the standard in the same participants. We then combined a model of the auditory periphery with generative EEG methods (temporal response functions, TRFs) to model the single-participant MMN and others ERPs responses to flat-intensity deviants, and used them to predict the effect of the same mechanism on time-varying looming and receding stimuli. By comparing actual vs predicted responses, we could investigate whether and how the ERPs responses to looming sounds is specific to their arousing or salient nature or, on the contrary, explained away by generic auditory mechanisms.

2. Methods

2.1. Participants

$N = 18$ volunteer participants (9 females, Median age = 25 years, Standard deviation $SD = 4.9$, all right-handed) took part in the experiment. All participants reported no neurologic or psychiatric diseases and normal hearing. The experimental protocol was approved by Institut Européen d'Administration des Affaires (INSEAD)'s Institutional Review Board (protocol ID: 2021–51), and all participants gave written informed consent before the start of the study. They were financially compensated for their participation (25 euros/participant). No part of the study procedures and analyses were pre-registered prior to the research being conducted. We report how we

determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

2.2. Stimuli

1000 Hz pure tones of different time-varying intensities were generated with custom Python software. Standard sounds had a duration of 300 ms, with constant (root-mean-square) RMS intensity. All three types of deviants (flat, looming and receding) had a duration of 600 ms (Fig. 1A). Flat deviants had the same constant RMS-intensity as standards. Looming deviants started at the same intensity as flat and standards, but their RMS-intensity increased linearly by 15 dB over the duration of 600 ms. Receding deviants started at the maximum intensity reached by looming deviants, and their RMS-intensity decreased linearly by 15 dB over the duration of 600 ms.

2.3. Procedure

We used an MMN *oddball paradigm* with standard (i.e., frequent and repetitive sounds) mixed with deviant (i.e., rare and unpredictable) sounds, the deviants being presented in a random order (Näätänen et al., 1993). The interval between the end of the sound and the beginning of the next one was set at 600 ms with a jitter of 50 ms. Standards represented 80% of all sounds, all deviants represented 20% (i.e., 6.7% of each), corresponding to 1602 standards and 133 deviants of each type. Three blocks of auditory stimuli were delivered for each participant, each block including 534 standards and 133 deviant sounds of all three types, with an inter-block interval duration of 5 min. Subjects were seated in a comfortable chair, in a quiet testing room and started to listen to the experiment. Sounds were presented binaurally over headphones, delivered by Python software. The participants were naive with respect to the hypotheses under test. Participants were

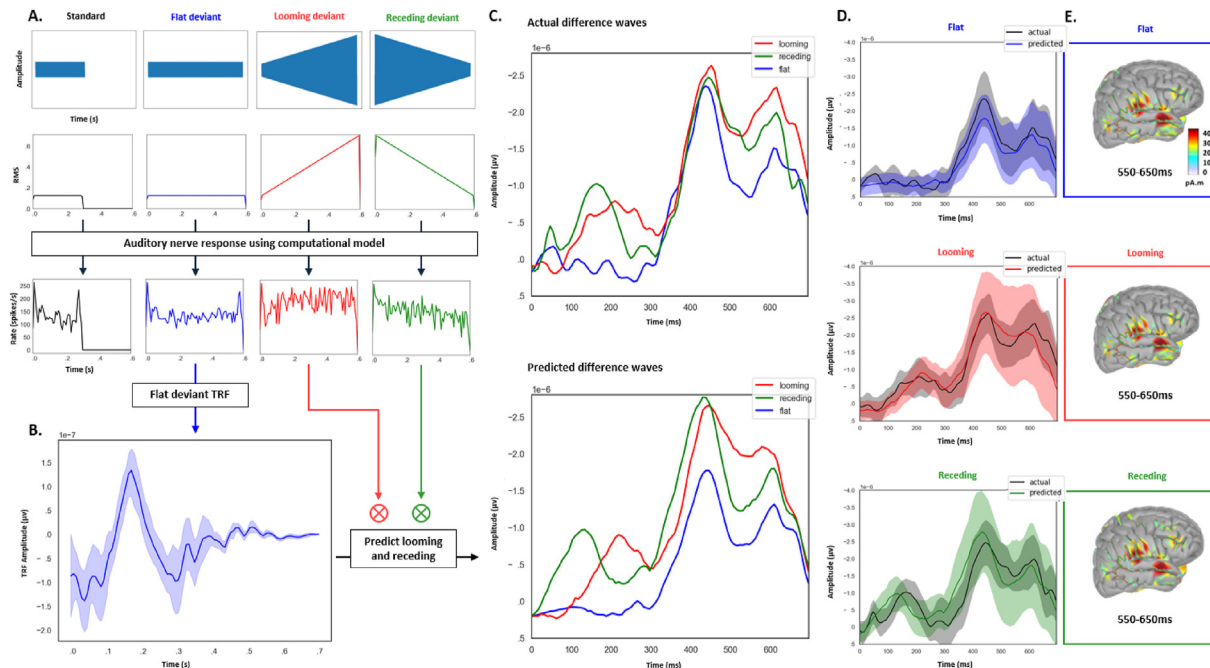


Fig. 1 – Cortical asymmetries between looming and receding sounds are explained away by the auditory periphery. A. We collected scalp EEG mismatch negativity (MMN) data in response to deviant stimuli that had either dynamic (looming and receding) or constant (flat) level differences to the standard, in the same participants. Top: waveform and RMS intensity profiles for all four types of stimuli. Bottom: simulated auditory nerve response for the four types of stimuli, using the computational model of Zilany et al. (2014). **B.** We then modeled the transfer function between the simulated auditory response and ERP difference wave for flat-intensity deviants (blue) using temporal response functions (TRFs), and used them to predict the effect of the same mechanism on other two stimuli (red: looming, green: receding). **C.** Predicted response to looming and receding sounds matched the observed difference waves almost perfectly (looming: 45%; receding: 33% explained variance). Top: Grand average of the observed difference waves (deviant minus standard) at the Fz sensor. Bottom: Predicted difference waves according to the flat TRF model. **D.** Cluster permutation test at the Fz sensor for observed vs predicted flat (top, blue), looming (middle, red), receding (bottom, green). No statistically significant difference was observed between the predicted responses and the observed responses for either type of stimuli. **E.** Sources localization in the right lateral cortical surface for observed flat (top), looming (middle) and receding (bottom) responses. All three types of deviants generated a right anterior temporal source (superior and middle temporal gyri) and a right posterior temporal source (auditory cortices and *planum temporale*), again with no statistical difference between stimuli. Taken together, these results evidence that MMN responses to looming and receding sounds result from the same cortical mechanism that generate MMN responses to unremarkable constant-level deviants.

asked not to pay particular attention to sounds, and to fix their attention on a relaxing mute video.

2.4. EEG recording

EEG signals were recorded using a 64-channel EEG (actiCHamp, Brain Products GmbH, Germany) with a sampling rate of 1000 Hz. Bandpass was set between .01 and 100 Hz. EEG sensors were placed according to the 10-10 system (Seeck et al., 2017) and Cz was set as the reference electrode. Sound onset triggers were sent to the EEG acquisition computer by a Cedrus StimTracker (Cedrus Corporation, San Pedro, CA) to control synchronization between the stimulus presentation and the appearance of the trigger on the recorded EEG.

2.5. Data pre-processing

Preprocessing and EEG analyses were performed with EEGLab/Matlab R2022b (Delorme & Makeig, 2004) and replicated with Python the minimum norm estimate MNE (Gramfort et al., 2013). First, continuous raw EEG data were low-pass filtered at 30 Hz and high-pass filtered at .1 Hz. A notch filter was applied (Parks–McClellan filter, 50 Hz). Second, these cleaned data were 1 Hz high-pass filtered to perform the independent component analysis (ICA). ICA was performed according to the recommended Makoto Miyakoshi EEGLab-pipeline (Makoto's preprocessing pipeline (n.d.). Retrieved April 23, 2020, from https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline). To apply ICA weights, continuous raw EEG data were again low-pass filtered at 30 Hz, high-pass filtered at .1 Hz with a 50 Hz notch filter. Bad electrodes (defined by an amplitude standard deviation $< 2 \mu\text{V}$ or $> 100 \mu\text{V}$) were interpolated with the trimOutlier EEGLab plug-in. Data were then cleaned with the clean_rawdata plug-in. ICA weights obtained at the first step were then applied to this EEG dataset. The IClab EEGLab plug-in was used to label independent components (among 7 labels: Brain, Muscle, Eye, Heart, Line noise, Channel noise and Other), and components reflecting eye artifacts were removed. EEG data were then re-referenced to the average reference and we segmented the EEG continuous data into epochs of 900 ms, ranging from -100 to 800 ms relative to sound onset. We applied a baseline correction for each trial before stimuli onset (-100 to 0 ms). All artifactual epochs, with voltage changes exceeding $\pm 100 \mu\text{V}$, were rejected from the analysis (Delorme & Makeig, 2004). Two subjects with more than 32% of epochs rejected were excluded from analysis, leaving 16 participants in the final analysis.

2.6. Event-related potentials analyses

ERPs of each participant were obtained by averaging separately each deviant and the standard stimuli using ERPlab software (Lopez-Calderon & Luck, 2014). Grand-averages were performed by averaging epochs for each condition in all participants. In this study, MMN responses were obtained from the difference waveform between deviants and standards (Näätänen et al., 1993), and compared between our three deviant conditions (looming, receding and flat). There is ambiguity in the community about what the term MMN refers to; for some authors, incl. classic studies by Näätänen et al., 1993,

2011a, 2011b, MMN is defined somewhat agnostically as the cluster of potentials identified in the difference wave of an oddball paradigm, regardless of whether they result from the subtraction of independently-evoked sensory activity (which may reflect e.g. neural adaptation in the standards), or to the specific cognitive process of registering change; others only equate MMN with the latter type of process (considered as “genuine” deviance detection; May & Tiitinen, 2010), and regard the former as a mere “subtraction artefact” (Fishman, 2014). In this work, we use the expression “MMN-like components” to refer to their former definition as components seen in compatible latency and spatial windows of the difference wave, without prescribing what underlying cortical mechanism may generate them. In particular, we do not imply that such components as “pure” deviance detection components, nor attempt to isolate such components experimentally by e.g. swapping standards and deviant stimuli in an inverse oddball block (see discussion for more details). For each ERP, peak latencies, peak amplitudes and area-under-curves (AUC) were automatically measured at the Fz sensor, using EEGLab software (Delorme & Makeig, 2004). For each ERP component, a mixed analysis of variances (ANOVA) with the within-subject factor stimulus type (the three types of deviants) was calculated. Post-hoc comparisons were made using independent sample t-test. Differences were considered significant when p value was $< .05$. JASP software (JASP Team (2022), version .16.3) was used for statistical analysis.

2.7. Auditory periphery modelling

In order to model the cortical response to looming, receding and flat deviant sounds, we used a computational model to simulate the effect of inner-ear-cell and auditory nerve (AN) non-linearities on the RMS profile of the three types of sounds. The model, described in Zilany et al. (2014) and implemented as a web application at https://urhear.urmc.rochester.edu/webapps/home/session.html?app=UR_EAR_2022a is one of two Auditory Nerve models to choose from: Zilany et al. (2014) and Bruce et al. (2018). We used the model developed by Zilany et al. (2014) with human parameters for sharpness tuning and middle ear filter model provided by Bruce et al. (2018). The Quick Plot functionality, which implements a single-CF Auditory Nerve model, was used to generate average discharge rate plots at a single characteristic frequency (CF) of 1000Hz for each of the three deviants.

2.8. TRF analysis

To simulate the extent to which auditory non-linearities could explain the differential response to time-varying looming and receding stimuli, we used generative EEG methods (temporal response functions, TRFs) to model single-participant ERPs responses to flat-intensity deviants, and used them to predict the effect of the same mechanism on time varying looming and receding stimuli. TRFs are impulse response functions that describe the relationship between the input and the output of a linear, time-invariant system. TRFs operate under the assumption that for a stimulus s at time t there exists a linear convolution with $s(t)$ that results in the output of the system at that time $r(t)$. For a system with N recording

channels, we can represent the neural response at time t and a specific channel n as the sum over all time lags τ of the linear convolution of stimulus characteristic $s(t)$ and the TRF for that particular channel $w(\tau, n)$.

$$r(t, n) = \sum_{\tau} w(\tau, n) s(t - \tau) + \varepsilon(t, n) \quad (1)$$

This linear convolution, referred to as the TRF, is estimated by minimizing the mean squared error between the actual and predicted responses using regularized ridge regression. In this study, we used the mTRFpy library (Crosse et al., 2016).

For each participant, we modelled the flat response using the following procedure: first, the AN response to flat sounds was subtracted by the AN response to standard sounds to get the stimulus difference (input). Second, all EEG epochs in which the flat deviant was presented were subtracted with the standard epoch preceding it to get the “flat-standard” EEG difference wave (output). We then estimated the input–output TRF by first, doing an exhaustive search for the best regularization parameter based on their cross-validated correlation between the predicted and measured response; and second, using the regularization parameter with greatest accuracy to build the final model. We obtained a single TRF for each participant. That TRF was then convoluted with the looming-standard and receding-standard stimuli to predict the neural response to the respective sounds (Fig. 2). We ran temporal cluster permutation tests to test for any significant differences between the actual and predicted responses.

2.9. Sources localization

The estimation of cortical current source density was performed with Brainstorm (Tadel et al., 2011). EEG electrode positions were aligned to the standard Montreal Neurological Institute (MNI) template brain provided in Brainstorm. The mean head model was computed with the OpenMEEG Boundary Element Method for all participants (Gramfort et al., 2010). A noise covariance matrix was computed for each participant by taking the 100 ms baseline period of each trial. For each subject, we computed one sensor-level average per condition (standard, looming, receding and flat). We then estimated sources for each average during the time window [−100; 800 ms] using different methods of standardization: minimum norm estimate, dynamical Statistical Parametric Mapping (dSPM; Dale et al., 2000) and standardized Low resolution brain Electromagnetic Tomography (sLORETA; Pascual-Marqui et al., 2002). Source cortical maps were then compared with permutation paired t-test between the different types of sounds, in the different time-windows identified as significant in the grand average, and in ROIs recognized as of interest for MMN (Alho, 1995).

2.10. Transparency and Openness (TOP) statement

All anonymized data and digital materials necessary and sufficient to reproduce analyses are accessible at <https://osf.io/wjsbf/>, <https://osf.io/5et62/> and <https://osf.io/qfv5y/>.

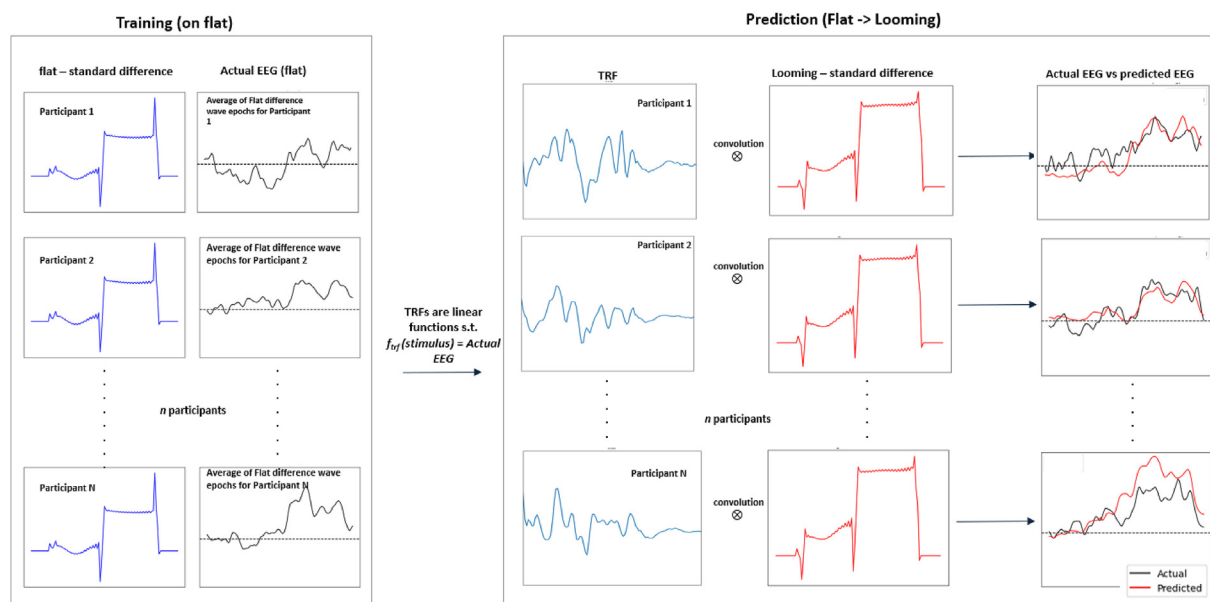


Fig. 2 – TRF analysis. For each participant, we modelled the flat response using the following procedure: first, the auditory nerve (AN) response to flat sounds was subtracted by the AN response to standard sounds to get the stimulus difference (input). Second, all EEG epochs in which the flat deviant was presented were subtracted with the standard epoch preceding it to get the “flat-standard” EEG difference wave (output). We then estimated the input–output TRF by first, doing an exhaustive search for the best regularization parameter based on their cross-validated correlation between the predicted and measured response; and second, using the regularization parameter with greatest accuracy to build the final model. We obtained a single TRF for each participant. This TRF was then convoluted with the looming-standard and receding-standard stimuli to predict the ERPs to these respective stimuli.

3. Results

3.1. Event related potentials

We recorded 64-channel EEG from a sample of $N = 18$ participants (9 females, median age = 25 years old) while they were presented a 30-min sequence of frequent pure tones (1000 Hz, 300 ms) with constant-intensity combined with rare deviants that were both longer (600 ms) and had either constant (flat) or dynamically changing intensity (looming and receding; Fig. 1A).

At the scalp level, difference-waves between deviant and standard sounds showed a striking succession of “MMN-like” negative deflections, which extended approximately 170 ms, 450 ms and 600 ms after stimulus onset (Fig. 1C). The first peak, which corresponded to a fronto-central midline distribution with maximum negativity at Fz (Fig. 1 Suppl. Material), was compatible with a MMN due to the initial dynamic difference between stimuli. Compared to receding, looming elicited a later difference-wave peak latency (206.5 vs 175.5 ms, $p = .022$) with a wider area under curve (AUC: .116 vs .053, $p = .002$) (Table 1). Expectedly, there was no such peak for flat deviants, which did not differ from standards in that time range. Because the initial intensity of looming sounds were set equal to both standards and flat deviants, this MMN-like difference-wave component was not trivially explained by intensity differences at the onset, but rather translates the specific temporal dynamics of the looming and receding stimuli. The second peak, elicited between 400 and 500 ms - thus about 150 ms after the end of standard sounds, was compatible with a response to the end of standard sounds and not with a MMN. It did not correspond to any visible component in the individual deviant waves (Fig. 2 Suppl. Material), and plausibly resulted from the cessation of the sustained potential evoked by the standard. Correspondingly, it occurred in a similar manner in all three types of deviants, with no statistical difference in either evoked responses characteristics (peak amplitudes, peak latencies and AUC) or surface amplitude maps. We do not further discuss it here.

Finally, all three types of deviants elicited a late negative component between 550 and 650 ms which, presented a fronto-central distribution with maximum negativity at Fz (Fig. 1 Suppl. Material). Contrary to the second component, components in the 550–650 ms time window were all clearly visible in the deviant waves, while the standard response remains relatively steady from 500 to 800 ms (Fig. 2 Suppl. Material). This late component also largely differentiated looming and receding sounds from flat sounds both in terms of peak amplitude, which was higher for looming ($-1.96 \mu\text{V}$) than flat ($-1.43 \mu\text{V}$, $p = .01$); peak latency, which was earlier for receding (599 ms) and looming (602 ms) than flat (618 ms, $p = .03$); and AUC, wider for looming (.263) than receding (.197, $p = .05$) and flat (.181, $p = .023$) (Fig. 1C and Table 1). In sum, at the scalp level, looming sounds displayed a clear pattern of amplification of MMN-like responses in the difference wave, with wider responses in the range 150–250 ms and earlier and more intense responses in the range 550–650 ms. This pattern of result was consistent with other examples of looming amplification both in behavior (Neuhoff et al., 1998 (Neuhoff, 1998)), brain imaging (Seifritz et al., 2002) and electrophysiology (Shestopalova et al., 2018).

3.2. TRF model

To tease apart bottom-up auditory components from more specific top-down contributions in the response to looming and receding sounds, we then used generative EEG methods (temporal response functions, TRFs) to measure the extent to which the latter can be predicted from the former. First, we used a computational model of the auditory periphery (Zilany et al., 2014) to simulate the nonlinearities (i.e. loudness compression, temporal integration, onset and offset amplification) observed at the level of the auditory nerve, and how these affect the root-mean-square (RMS) profile of flat, looming and receding stimuli (Fig. 1A). We then used TRFs to model each participant's ERPs to the output of this auditory model for flat-intensity deviants. TRFs provide an approximation of the mapping between incoming stimuli (here, the difference wave between the peripheral encoding of flat deviants minus

Table 1 – Comparison of peak amplitudes, peak latencies and area under curves (AUC) for the three “early” “second” and “late” components, according to the looming, receding and flat conditions.

Conditions	looming	receding	Flat	Looming vs receding mean difference	Looming vs flat mean difference	Receding vs flat mean difference
Early component (150–250 ms)						
Peak amplitude (μV , SD)	1.24 (.46)	1.40 (.98)	.21 (.44)	.16 $p = .4$	NA	NA
Peak latency (ms, SD)	206.5 (29)	175.5 (28)	NA	31^a $p = .022$	NA	NA
AUC (SD)	.116 (.063)	.053 (.063)	NA	.063^a $p = .002$	NA	NA
Second component (400–500 ms)						
Peak amplitude (μV , SD)	2.55 (.98)	2.50 (1.3)	2.59 (1.4)	.05	.04	.09
Peak latency (ms, SD)	442.8 (18)	445.4 (16)	440.8 (17)	2.6	2	4.6
AUC (SD)	.315 (.17)	.308 (.21)	.285 (.22)	.007	.03	.023
Late component (550–650 ms)						
Peak amplitude (μV , SD)	1.96 (1.2)	1.61 (1.2)	1.43 (1.33)	.35	.53^a $p = .01$.18
Peak latency (ms, SD)	602 (24)	599 (22.4)	618 (17.6)	.03	16^a $p = .03$	19^a $p = .03$
AUC (SD)	.263 (.21)	.197 (.19)	.181 (.23)	.066^a $p = .05$.082^a $p = .023$.016

Measures of peak amplitudes, peak latencies and AUC were performed at the level of Fz sensor.

^a indicate the significant difference with the p value, all other statistical comparison being not statistically significant. NA: not assessable; ms: milliseconds.

standards) and the output EEG using a simple linear, time-invariant model represented by an impulse response (Crosse et al., 2016). Here, we trained a separate TRF for each individual participant that predicts the participant's ERPs to flat deviants (Fig. 1B). The predicted response closely matched the observed response (80% explained variance; Fig. 1D), and thus provided an accurate model of the cortical response to a simple change of stimulus duration, at constant level. Finally, we used that flat-deviant model to simulate the extent to which bottom-up auditory mechanisms could explain the response to time-varying looming and receding stimuli (Fig. 2). To do so, for each participant, we convoluted the TRF trained on their flat response with the non-linear peripheral encoding of the two other types of deviants (looming and receding). Strikingly, both predicted responses almost perfectly matched the observed response (Fig. 1D): in particular, predicted looming responses exhibited the same pattern than seen in actual responses, i.e., a later/wider peak in the 150–250 ms range, and an earlier/larger peak in the 550–650 ms range. We tested for statistical differences between the predicted and observed responses across participants using cluster-based permutation test, and neither was significant. The flat model explained 45% variance of the looming response, and 33% of the receding response. These figures are in the higher end of encoding accuracies usually reported for EEG TRF studies (Bednar & Lalor, 2020; de Cheveigné et al., 2018).

3.3. Sources localization

Finally, we extracted cortical current source densities in the 150–250 ms and 550–650 ms time windows, and checked for significant differences between the three deviants. All three types of deviants generated a right anterior temporal source (superior and middle temporal gyri), a right posterior temporal source (auditory cortices and *planum temporale*) and to a lesser extent a right prefrontal cortex source (Fig. 1E). We also highlighted a left inferior temporal gyrus source (Fig. 3 Suppl. Material). None of them differed statistically, suggesting that they were generated by a close mechanism (Fig. 4 Suppl. Material).

4. Discussion

Our results show that the large differences between cortical responses to looming and receding sounds exhibited at the scalp-level are in fact almost entirely explained away by non-linear encoding at the level of the auditory periphery. These cortical responses to looming and receding sounds result from the same cortical mechanism that generate MMN responses (i.e., at 150–250 ms) and later evoked potentials (i.e., at 550–650 ms) to unremarkable flat-level deviants. In other words, there is nothing cortex specific in the processing of looming sounds up to the level of the MMN response: all differences observed in MMN are the sole consequence of their particular physical morphology getting amplified and integrated by peripheral auditory mechanisms.

The sharp contrast seen here between the visually-salient scalp-level effects and their unassuming explanation by early peripheral differences should provide a sobering reminder

that not all cortical effects, even in relatively late time-windows such as seen here, proceed from top-down modulations by high-level decision variables, such as a stimulus' supposed physical, affective or social relevance. Rather, a lot of such differential amplification observed at the level of the cortex is in fact performed early and efficiently by feed-forward peripheral mechanisms that evolved precisely for the purpose of sparing subsequent networks with the necessity to implement such mechanisms. Our results are congruent with some other ERPs and MMN studies, highlighting that MMN may involve automatic activation of low-level feature detectors (or early stages of auditory processing) rather than higher level attention-dependent processing (Bishop et al., 2005). Our results also suggested that when a target has a greater magnitude of activation (here, looming and receding compared to flat sounds), it is detected more easily as previously demonstrated (Cusack & Carlyon, 2003). This perceptual asymmetry effect (observed here with sounds intensity modification) is also largely observed with speech sounds, MMN being sensitive to phonological and phonetic contrasts (Højlund et al., 2018; Scharinger et al., 2010; Scharinger et al., 2012). This effect was also observed in the visual system. Indeed, human subjects detect a change more easily when a feature is added than when a feature is deleted, suggesting that these perceptual asymmetries concern several sensory functions (Treisman & Gormican, 1988; Cusack & Carlyon, 2003; Treisman & Gormican, 1988).

Our study presents some strengths. First, this is the first study that used a generative EEG methods TRFs to model single-participant evoked potentials responses to flat-intensity deviants, and used them to predict the effect of the same mechanism on time varying looming and receding stimuli. Some studies have already assessed looming or receding in an MMN oddball paradigm and have highlighted some differences in amplitudes between conditions (Shestopalova et al., 2018), but none have attempted to differentiate the sensory vs. cognitive mechanisms for the integration of these stimuli. We also assessed the cortical current source densities for each type of deviants, and found consistent results with results of the TRF model.

Our study also presented some limitations. First, we only used a linear TRF model. Because this convolution technique is apt for linear systems, and as neural MMN system is probably highly non-linear, we could not exclude that this TRF model was not perfectly appropriate. Because we used a computational model to simulate the effect of inner-ear-cell and auditory nerve (AN) non-linearities on the RMS profile of the three types of sounds, it is possible that this procedure allowed us to simulate non-linearities that are entailed not only by the auditory periphery, but also by the cortical MMN system. Future work could attempt to model specifically-cortical non-linearities in a data-driven way, as done e.g. in Deneux et al. (2016). Second, in all our analyses, we identified MMN-like components in the simple difference wave between standard and deviant sounds, but did not attempt to control empirically for whether such components correspond to “genuine” deviance detection, e.g. by swapping deviant and standard stimuli. Thus, We only used the oddball paradigm described by Näätänen et al. to elicit an MMN, and we did not swapped the roles of deviant and standard sounds in a reversed oddball

block (see e.g. Althen et al., 2011). Doing so, we cannot rule out that some of the components evidenced here result from confounding physical differences between stimuli (as is trivially the case for the component seen at 450 ms) and, more generally, to lower-level explanations, such as differential N1 adaptation. Crucially though, such explanations still require accounting for the subtle ways in which the complexity of temporal stimuli translate into variations in the amplitude, latency, and morphology of ERPs produced by auditory generators – all mechanisms which, in fact, implicitly represent an index of the salience of stimulus representations in the brain (Fishman, 2014). Thus, regardless of their underlying mechanisms (explicit deviance detection, or encoding of stimulus features which are implicitly sensitive to salience), the early and late MMN-like components studied here reflect a cortical index that is sensitive to the specific dynamics of looming or receding stimuli, and which is associated with feed-forward peripheral (bottom-up) mechanisms rather than top-down ones.

Open practices

The study in this article has earned Open Data and Open Materials badges for transparent practices. The data and materials are available at: <https://osf.io/2ztwv/>.

CRediT authorship contribution statement

Sarah Benghanem: Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **Rudradeep Guha:** Formal analysis, Writing – original draft, Writing – review & editing. **Estelle Pruvost-Robieux:** Data curation, Formal analysis, Writing – review & editing. **Julie Lévi-Strauss:** Data curation. **Coralie Joucla:** Data curation, Formal analysis. **Alain Cariou:** Supervision, Validation, Writing – review & editing. **Martine Gavaret:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Jean-Julien Aucouturier:** Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – review & editing.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2024.05.018>.

REFERENCES

- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing*, 16, 38–51.
- Alho, K., Rinne, T., Herron, T. J., & Woods, D. L. (2014). Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies. *Hearing Research*, 307, 29–41.
- Althen, H., Grimm, S., & Escera, C. (2011). Fast detection of unexpected sound intensity decrements as revealed by human evoked potentials. *Plos One*, 6, Article e28522.
- Altmann, C. F., Hiraumi, H., Terada, S., Adachi, T., Votinov, M., Ono, K., Mima, T., & Fukuyama, H. (2013). Preattentive processing of horizontal motion, radial motion, and intensity changes of sounds. *Neuroreport*, 24, 861–865.
- Bach, D. R., Neuhoff, J. G., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology*, 74, 28–33.
- Bach, D. R., Schächinger, H., Neuhoff, J. G., Esposito, F., Salle, F. D., Lehmann, C., Herdener, M., Scheffler, K., & Seifritz, E. (2008). Rising sound intensity: An intrinsic warning cue activating the amygdala. *Cerebral Cortex*, 18, 145–150.
- Bednar, A., & Lalor, E. C. (2020). Where is the cocktail party? Decoding locations of attended and unattended moving sound sources using EEG. *Neuroimage*, 205, Article 116283.
- Bishop, D. V. M., O'reilly, J., & McArthur, G. M. (2005). Electrophysiological evidence implicates automatic low-level feature detectors in perceptual asymmetry. *Brain Res Cogn Brain Res*, 24, 177–179.
- Bruce, I. C., Erfani, Y., & Zilany, M. S. A. (2018). A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites. *Hearing Research*, 360, 40–54.
- Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: A MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in Human Neuroscience*, 10, 604.
- Cusack, R., & Carlyon, R. P. (2003). Perceptual asymmetries in audition. *Journal of Experimental Psychology. Human Perception and Performance*, 29, 713–725.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26, 55–67.
- de Cheveigné, A., Wong, D. D. E., Liberto, G. M. D., Hjortkjær, J., Slaney, M., & Lalor, E. (2018). Decoding the auditory brain with canonical component analysis. *Neuroimage*, 172, 206–216.
- Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21.
- Deneux, T., Kempf, A., Daret, A., Ponsot, E., & Bathellier, B. (2016). Temporal asymmetries in auditory coding and perception reflect multi-layered nonlinearities. *Nature Communications*, 7, Article 12682.
- Fishman, Y. I. (2014). The mechanisms and meaning of the mismatch negativity. *Brain Topography*, 27, 500–526.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7, 267.
- Gramfort, A., Papadopoulos, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: Opensource software for quasistatic bioelectromagnetics. *Biomedical Engineering Online*, 9, 45.
- Højlund, A., Gebauer, L., McGregor, W. B., & Wallentin, M. (2018). Context and perceptual asymmetry effects on the mismatch negativity (MMNm) to speech sounds: An MEG study.
- Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception & Psychophysics*, 78, 373–395.
- Leo, F., Romei, V., Freeman, E., Ladavas, E., & Driver, J. (2011). Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Experimental Brain Research*, 213, 193–201.

- Lopez-Calderon, J., & Luck, S. J. (2014). Erplab: An open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci*, 8, 213.
- May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47, 66–122.
- Näätänen, R., Kujala, T., Kreegipuu, K., Carlson, S., Escera, C., Baldeweg, T., & Ponton, C. (2011a). The mismatch negativity: An index of cognitive decline in neuropsychiatric and neurological diseases and in ageing. *Brain: a Journal of Neurology*, 134, 3435–3453.
- Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*, 48, 4–22.
- Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*, 30, 436–450.
- Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, 395, 123–124.
- Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): A review. *Methods Find Exp Clin Pharmacol*, 24(Suppl C), 91–95.
- Ponsot, E., Susini, P., & Meunier, S. (2015). A robust asymmetry in loudness between rising- and falling-intensity tones. *Attention, Perception & Psychophysics*, 77, 907–920.
- Ponsot, E., Susini, P., Saint Pierre, G., & Meunier, S. (2013). Temporal loudness weights for sounds with increasing and decreasing intensity profiles. *The Journal of the Acoustical Society of America*, 134, EL321–EL326.
- Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research*, 1077, 135–143.
- Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology: CB*, 19, 1799–1805.
- Rosenblum, L. D., Carello, C., & Pastore, R. E. (1987). Relative effectiveness of three stimulus variables for locating a moving sound source. *Perception*, 16, 175–186.
- Scharinger, M., Bendixen, A., Trujillo-Barreto, N. J., & Obleser, J. (2012). A sparse neural code for some speech sounds but not for others. *Plos One*, 7, Article e40953.
- Scharinger, M., Lahiri, A., & Eulitz, C. (2010). Mismatch negativity effects of alternating vowels in morphologically complex word forms. *Journal of Neurolinguistics*, 23, 383–399.
- Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: Effects of modality, trajectory, and gender. *Journal of Experimental Psychology. Human Perception and Performance*, 16, 303–316.
- Seeck, M., Koessler, L., Bast, T., Leijten, F., Michel, C., Baumgartner, C., & Beniczky, S. (2017). The standardized EEG electrode array of the IFCN. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 128, 2070–2077.
- Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., Elefante, R., & Di Salle, F. (2002). Neural processing of auditory looming in the human brain. *Current Biology*, 12, 2147–2151.
- Shestopalova, L. B., Petropavlovskaja, E. A., Semenova, V. V., & Nikitin, N. I. (2018). Mismatch negativity and psychophysical detection of rising and falling intensity sounds. *Biological Psychology*, 133, 99–111.
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, 2011, Article 879716.
- Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: The emotional impact of approaching and receding sound sources. *Emotion*, 10, 216–229.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48.
- Zilany, M. S. A., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *The Journal of the Acoustical Society of America*, 135, 283–286.