Non-linear relationships between the auditory mismatch responses and the level of inharmonicity of complex sounds

Aniela Brzezińska*[1], Bartosz Witkowski*[1], Tomasz Domżalski[1], Małgorzata Basińska[2], Krzysztof Basiński[1]

1. Auditory Neuroscience Laboratory, Department of Psychology, Medical University of Gdańsk, Gdańsk, Poland
2. Division of Quality of Life Research, Department of Psychology, Medical University of Gdańsk, Gdańsk, Poland

Abstract

Predictive processing accounts of perception suggest that the brain is using an internal generative model to form predictions about the incoming stimuli. Precision weighting is an important facet of predictive processing theories and has been extensively studied with mismatch responses such as mismatch negativity (MMN). Harmonicity is a feature of sound that is important for auditory perception and previous research has shown that it modulates the brain's mismatch responses. Since inharmonic sound spectra contain more information (have higher information entropy), inharmonicity has been suggested to be involved in precision weighting. In this study we explored this issue by parametrically modulating the level of inharmonicity applied to synthetic sounds and recording MMN and P3a responses from healthy volunteers. Our results show that a sigmoid function models the relationship between inharmonicity and MMN amplitude better than any linear or polynomial function. Furthermore, P3a amplitude has an inverted-U relationship with inharmonicity and peaks at inharmonicity levels just below the threshold for pitch discrimination. These results are consistent with the hypothesis that inharmonicity impairs F0 extraction above a certain threshold and does not serve as an index of precision in the auditory system.

Introduction

If a complex tone consists of frequencies that are integer multiples of one fundamental frequency (F0), it is considered harmonic. Conversely, complex tones that include non-integer multiples of one frequency are considered inharmonic. Harmonicity is important in different domains of auditory perception. Many ecologically salient sounds have harmonic spectra - these include human speech, animal vocalizations, or sounds of musical instruments. Inharmonic sounds are harder to detect in noise, suggesting a role in auditory scene analysis (de Cheveigné, 2021; McPherson et al., 2022; Popham et al., 2018). Pitch discrimination tasks that rely on F0 extraction are severely impaired for inharmonic sounds (McPherson & McDermott,

2018). Finally, harmonic sounds may be more efficiently encoded in memory because their spectra can be easily compressed using the F0. This idea is supported by results showing that pitch discrimination across time delays is impaired for inharmonic sounds (McPherson & McDermott, 2020).

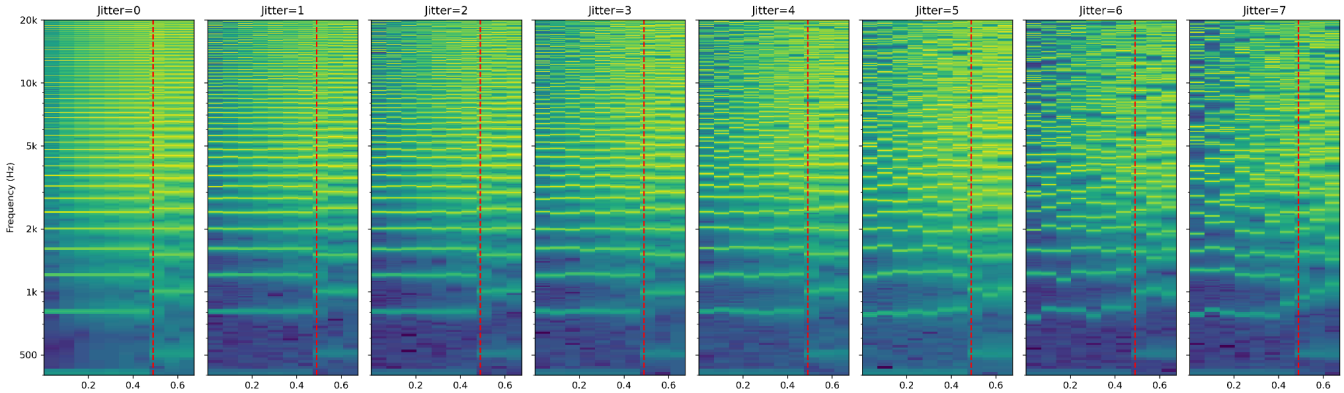
Predictive processing accounts of perception suggest that the brain is using an internal generative model to form predictions about the incoming stimuli (Aizenbud et al., 2025; Bastos et al., 2012; Friston, 2005; Keller & Mrsic-Flogel, 2018; Mumford, 1992; Rao, 2024; Rao & Ballard, 1999). Whenever these predictions are violated, a prediction error response is propagated throughout the cortical hierarchy and used to change the contents of the generative model. Importantly, prediction errors are suggested to be weighted by the precision (or uncertainty) of the context such that generative model update is more likely for highly precise prediction errors (Kanai et al., 2015; Palmer et al., 2019; Yon & Frith, 2021). Precision weighting is an important facet of predictive processing theories and has been extensively studied in recent years (Bowman et al., 2023; Milne et al., 2021; Zhao et al., 2025). In human neuroimaging, this process can be studied by employing auditory oddball paradigms and recording event-related potentials (ERPs) using electroencephalography (EEG) (Basiński et al., 2023; Heilbron & Chait, 2018). The mismatch negativity (MMN) is a widely studied ERP component that is established as an electrophysiological trace of precision-weighted prediction errors in the cortical auditory system (Friston, 2005; Garrido et al., 2009; Winkler & Czigler, 2012). MMN is often followed by P3a, a positive ERP component associated with an automatic attention shift to the perceived oddball (Polich & Criado, 2006).

The spectrum of a harmonic sound can be reliably represented by the harmonic series starting at the F0. Inharmonic sounds have spectra that include mathematically unrelated frequencies. Thus, inharmonic sounds carry more information (they have higher information entropy) than harmonic sounds. This observation, in conjunction with psychophysical results suggesting impaired perception of inharmonic sounds (McPherson et al., 2022; McPherson & McDermott, 2018, 2020), has lead to the hypothesis that harmonicity can potentially serve as an index of precision for auditory stimuli (Basiński et al., 2023; Quiroga-Martinez et al., 2022). A recent study explored this idea empirically by exposing participants to harmonic and inharmonic sounds in a roving oddball paradigm (Basiński et al., 2025). Using synthetic stimuli, the authors induced inharmonicity to the harmonic complex tone by introducing a random jitter to all frequencies above the F0. Surprisingly, inharmonic sounds with a constant jittering pattern throughout the sequence generated similar MMN and larger P3a responses than harmonic sounds. However, when the jittering pattern was randomized for each consecutive sound, both MMN and P3a responses were not detectable. This finding is consistent with precision weighting, as inharmonicity leads to a decrease in mismatch responses, at least for sequentially changing sounds. However, since the study compared harmonic to heavily jittered inharmonic sounds, it is unclear if the effect is caused by precision weighting (an increase in the uncertainty of predictions) or a fluctuating perception of the standard's pitch caused by constant jittering of harmonics (similar to the "many standards" paradigm (Carbajal & Malmierca, 2018)).

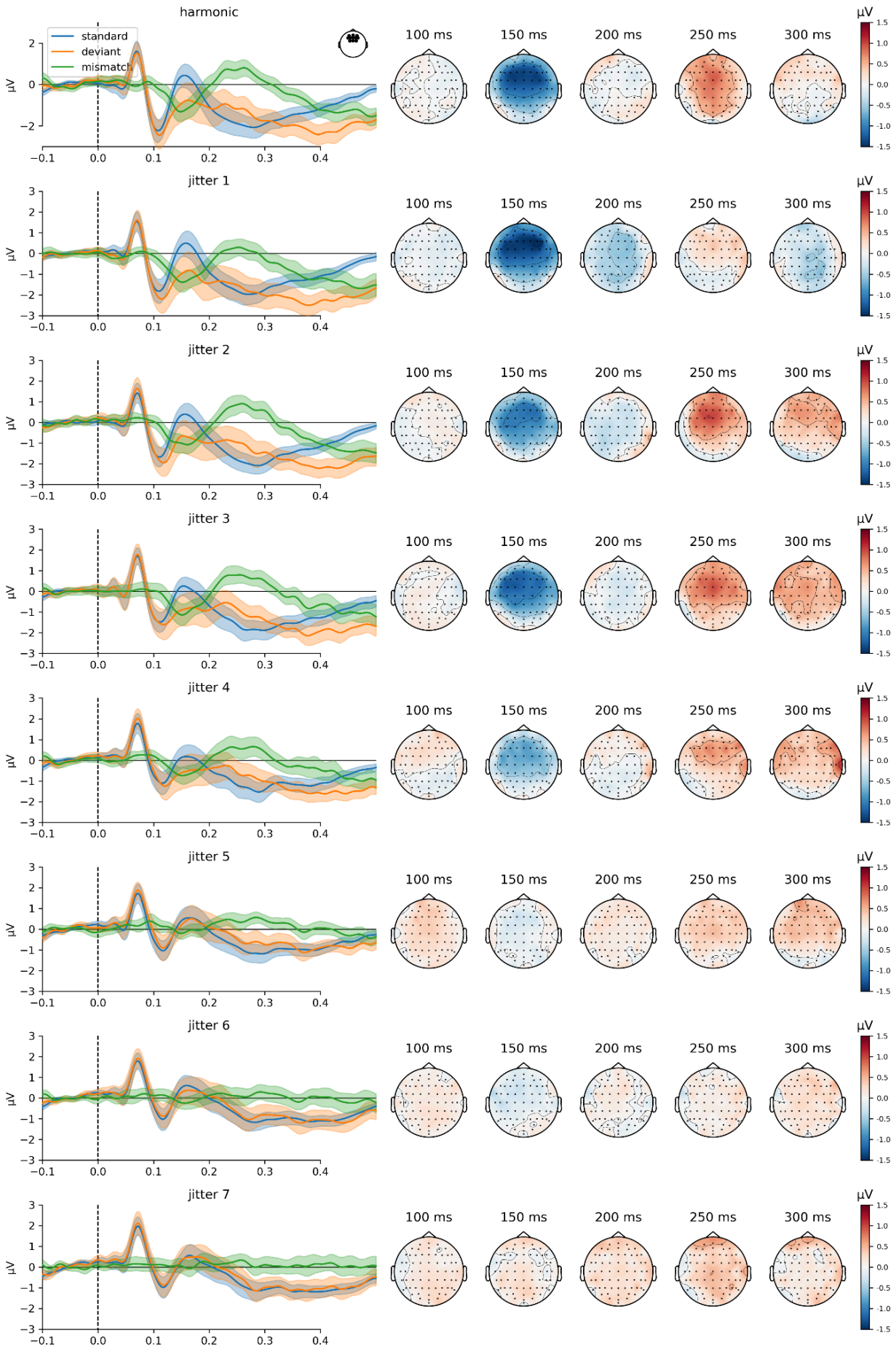
In the current study, we aimed to explore the relationship between the amount of inharmonicity and mismatch responses further. We prepared synthetic sound sequences with randomly jittered frequencies and systematically manipulated the rate of this jitter. We recorded MMN and P3a mismatch responses to pitch deviants in these sequences using EEG. We investigated two alternative hypotheses. Since the increase in the level of inharmonicity directly corresponds to an increase in information entropy of sounds, the precision-weighted prediction errors to increasingly inharmonic sounds theoretically should manifest in a linear decrease of MMN amplitudes. Alternatively, the auditory system may employ F0 extraction to detect deviants. In this case, we expect the MMN amplitudes to remain similar for small jitter rates, as F0 extraction would not be impaired by low-to-moderate levels of inharmonicity. Once a threshold value is reached, we expect mismatch responses to drop abruptly to near-zero levels (undetectable MMN) for larger jitters, presenting as a sigmoid relationship. This is because if the F0 extraction is impaired, the perception of the standard becomes unstable and any incoming pitch deviants would not elicit a mismatch response.

Our results show that a sigmoid function models the relationship between inharmonicity and MMN amplitude data better than any linear or polynomial function. This is consistent with the hypothesis that inharmonicity impairs F0 extraction above a certain threshold and does not serve as an index of precision in the auditory system. These findings are further supported by behavioral data, showing that pitch discrimination is impaired at a similar inharmonicity threshold. Finally, we also observed a slight peak in the P3a responses around the inharmonicity threshold which can reflect the higher computational demand associated with detecting pitch deviants at the threshold level.

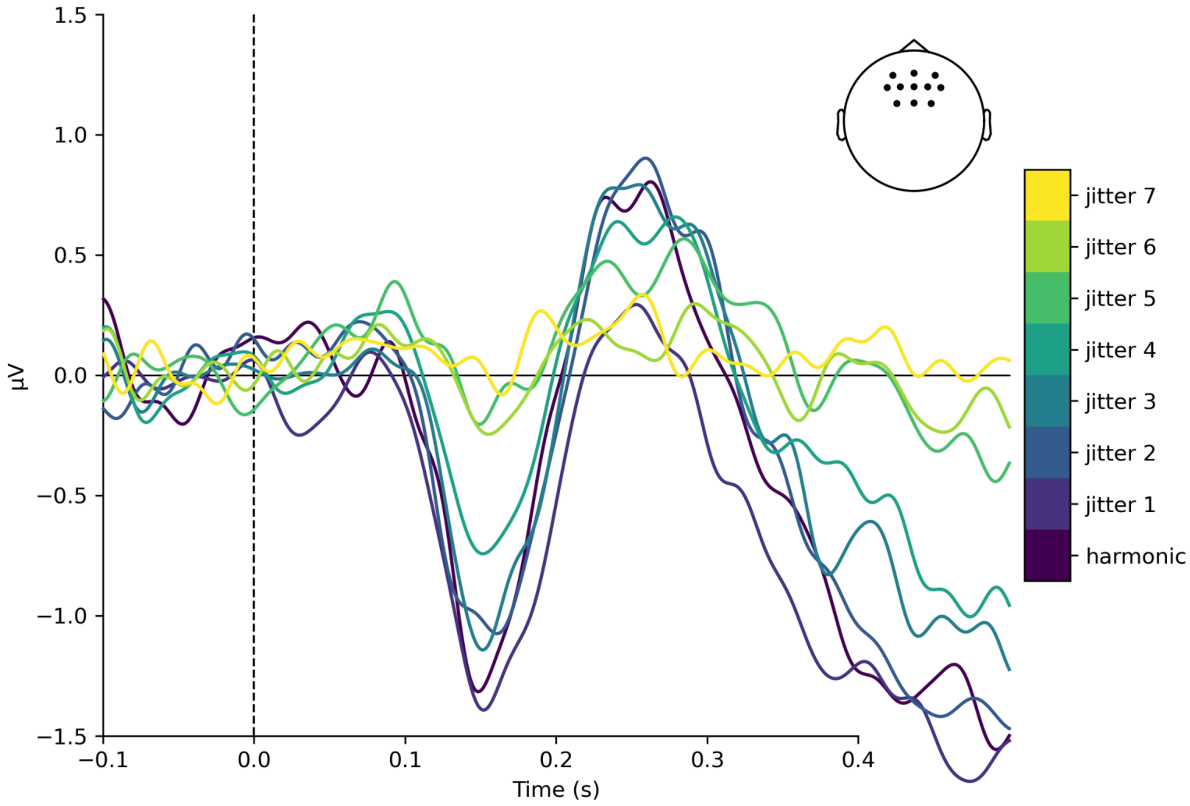
Results



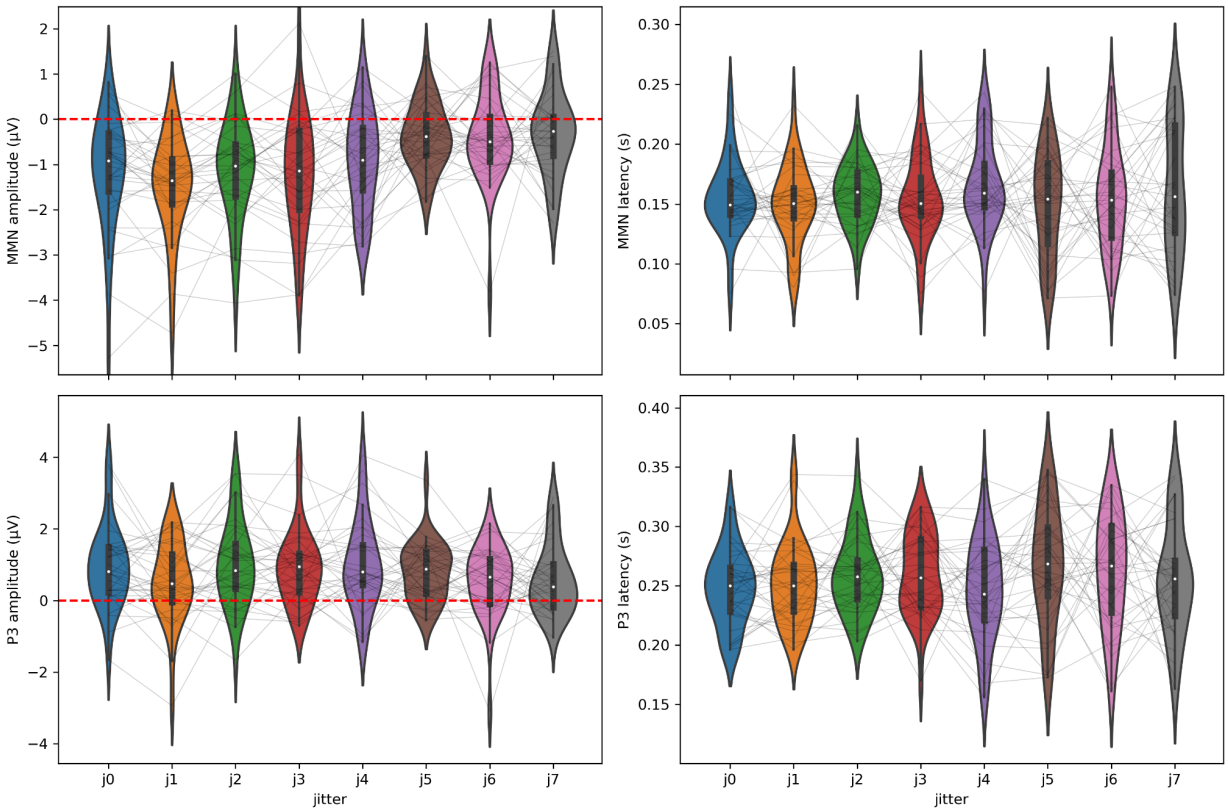
[Figure 1. Spectrograms of sound sequences used in the roving oddball paradigm. Each panel displays the spectral structure of a sound sequence used in the roving oddball paradigm for a given jitter condition. The x-axis represents time, and the y-axis represents frequency in Hz. For clarity of presentation, the silence between the sounds was removed, so the time axis does not reflect the actual ISI (600 ms). Red vertical lines mark the point where the F0 changes, introducing a pitch deviance. As jitter increases, the spectral structure of the sounds becomes increasingly irregular.]


harmonic
standard
deviant
mismatch
jitter 1
jitter 2
jitter 3
jitter 4
jitter 5
jitter 6
jitter 7
µV
100 ms
150 ms
200 ms
250 ms
300 ms

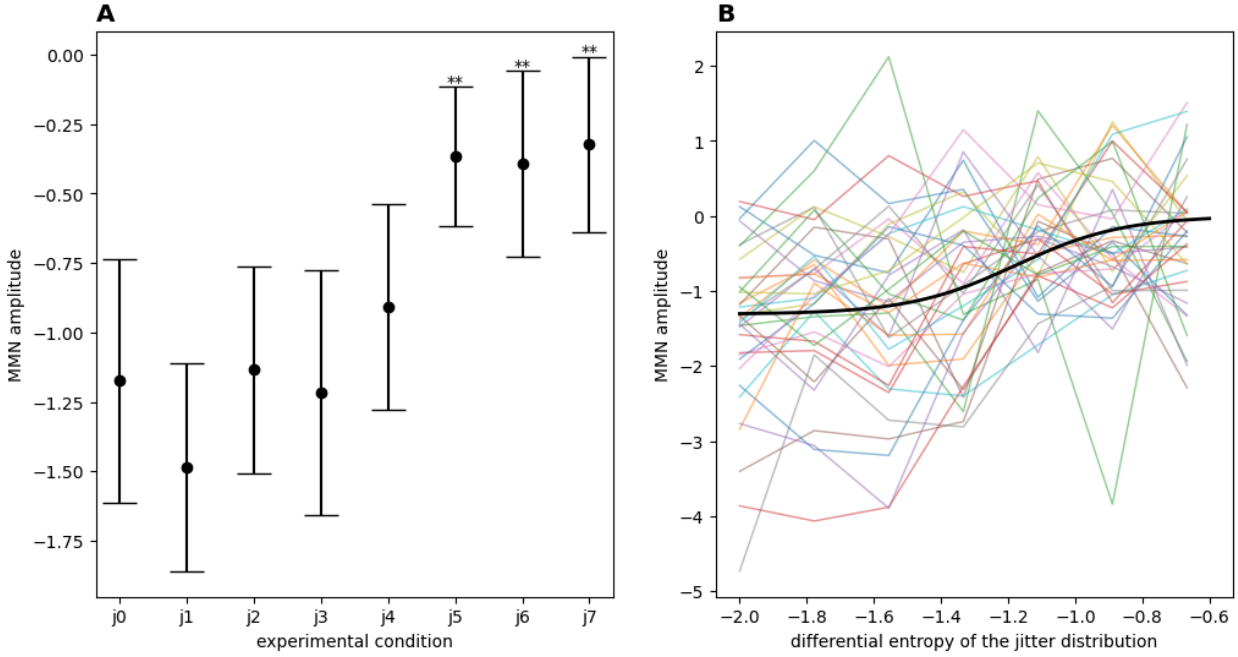
[Figure 2. Event-related potentials (left) and their topographies (right) for harmonic (first row) and seven inharmonic conditions. Event-related potential traces show grand-average responses to standards (blue), deviants (orange), and the difference waves (green) for fronto-central channels (Fz, F1, F2, AF3, AF4, AFz, F3, F4, FC1, FC2, FCz). Color-shaded areas represent 95% confidence intervals. Scalp topographies present signal strength at chosen latencies.]



[**Figure 3. Mismatch wave differences between experimental conditions.** Each trace is a difference wave between deviant and standard stimuli in a given jitter condition. Confidence intervals were not shown for visual clarity but can be seen in Figure 2.]



[**Figure 4. Peak measures differences between conditions.** Panels show MMN mean amplitude (upper-left), P3a mean amplitude (lower-left), MMN peak latency (upper-right), and P3a peak latency (lower-right). Violin plots show distributions of obtained results. Grey lines connect observations from the same participant.]



[**Figure 5. MMN amplitude as a function of experimental condition. (A)** Mean MMN amplitudes in experimental conditions. Stars indicate significant difference from condition “j0” (harmonic sounds). **(B)** Sigmoid curve (black line) as a function of the differential entropy of jitter factor distribution. Note that the ticked values on the x axis correspond to experimental inharmonic conditions j1-j7. Colored lines connect observations from the same participant.]

Procedure outline

In order to measure the MMN and P3a mismatch responses, we performed an EEG experiment using a version of the roving oddball paradigm (Cowan et al., 1993; Garrido et al., 2008). This paradigm typically involves presenting a sequence of repeating sounds at a specific F0, followed by another sequence at a different, randomly chosen F0, followed by another sequence and so on. The first sound after the F0 shift generates a mismatch response, but it becomes the new standard after several repetitions. The roving paradigm ensures that standards and deviants have the same acoustical properties, with mismatch induced by the context of the sequence. In this study, we systematically manipulated the level of inharmonicity between the eight experimental conditions. In the condition *j0* we used harmonic complex tones. For conditions *j1 - j7*, we applied random jittering to all frequencies above the F0, for each sound in the sequence. The jitter values were drawn from increasingly broader uniform distributions. This allowed us to systematically test the mismatch responses to progressively higher amounts of inharmonicity.

Mismatch negativity analysis

In the EEG experiment, mean MMN amplitudes differed significantly from zero for all conditions (means and t-test results are reported in Supplementary Table 1). ANOVA showed that harmonicity affected MMN amplitudes ($F(7, 238) = 9.12$, $p_{Greenhouse–Geisser} < 0.001$). To explore pairwise differences between jitter conditions, post-hoc comparisons were performed. MMN amplitude differed significantly between lower and higher levels of inharmonicity. Specifically, jitter levels *j0*, *j1*, *j2*, and *j3* differed significantly from j5, j6, and j7, (with p-values ranging from $p = .0044$ to $p < .0001$; see ST2), indicating a significant reduction in MMN amplitude at higher jitter levels. However, pairwise comparisons between all low jitter levels (j0–j4) and between all higher jitter levels (j5–j7) were not statistically significant (all p-values > .05), suggesting that the most pronounced changes in MMN occurred between the j4 and j5, rather than incrementally between all adjacent levels (Figure 3, ST3). These results suggest an abrupt decrease in the strength of mismatch responses at the threshold value rather than a gradual change, as in precision weighting.

To further explore how MMN amplitude depends on sound harmonicity, we used linear, polynomial and non-linear mixed effects models with differential entropy of jitter distribution as the independent variable. Since the jitters were drawn from a uniform distribution, with support ranging from $-j_c$ to j_c, where j_c was condition specific jitter rate, differential entropy I was expressed as:

$$I = log_{10} 2j_c.$$

Of note, differential entropy is not an exact continuous analogue of discrete entropy and does not share all its properties. Differential entropy can assume negative values, as was the case in our experiment, and zero differential entropy does not indicate certainty. Nevertheless, higher values of differential entropy indicate more uncertainty. In our experiment, higher entropies corresponded to more inharmonic sounds. For an outcome that is certain, differential entropy is not defined so in the further analyses we omit the harmonic condition with zero jitter rate.

We conjectured that the relation between inharmonicity and MMN amplitude is sigmoidal rather than linear and can be modeled as the following function:

$$MMN_{amplitude} = \frac{L}{1+e^{k(I-I_0)}},$$

where I is differential entropy of jitter sampling distribution, I_0 is the I value at the inflection point, k determines the slope at the inflection point and *L* is the left asymptote of the MMN amplitude (right asymptote is at 0, reflecting the disappearance of MMN for highly inharmonic sounds). Likelihood ratio test showed that the sigmoid mixed effects model with random left asymptote showed significantly better fit than the null model ($\chi^2(2) = 75.48$, $p<0.001$). Linear model with random intercept also showed significantly better fit to the data than the null model ($\chi^2(1) = 46.90$, $p < 0.001$). Adding higher degree parameters to the model did not significantly improve model fit (see Supplementary Table 4 for details). At the same time, both the Akaike and Bayesian information criteria were markedly better for the sigmoid model (AIC = 664.6, BIC = 682.2) than for the linear model (AIC = 691.2, BIC = 705.2), which suggests that the relationship between inharmonicity and MMN amplitude is indeed better modelled by a sigmoid than a linear or polynomial function. This corroborates the results of the contrast analysis. Since the linear and sigmoid models were not nested, we refrain from reporting the *p* value for the likelihood ratio test for the comparison of these two models, as this comparison would not yield a meaningful *p* value.

The I_0 parameter of the fitted curve was -1.17, suggesting that the MMN decays rapidly for jitter rate around 0.033, that is between *j4* and *j5*, convergent with the results of the contrasts analysis. The fitted curve is presented in Figure 5B. The details of both sigmoidal and linear models are reported in the Supplementary Table 5.

MMN peak latency was not significantly related to sound harmonicity, both in ANOVA ($F(7, 238) = 0.82$, $p_{Greenhouse-Geisser} = .538$, see also ST2) as well as in linear, quadratic and cubic mixed models (all p-values > .05, see also ST4 and ST5).

P3a analysis

P3 amplitudes (Figure 4C) differed significantly from zero for all harmonicity conditions (ST1 for details). ANOVA results did not indicate any significant effect of condition ($F(7, 238) = 1.78$, $p_{Greenhouse–Geisser} = 0.113$, see also ST2). However, we observed a significant relationship between the differential entropy of the jitter rate distribution and P3 amplitude in the mixed models analysis. The quadratic model had a significantly better fit than both linear ($\chi^2(2) = 6.93$, $p = 0.031$) and null models. The quadratic function had a shape of an inverted parabola with a predicted maximum at entropy = -1.41, corresponding to jitter conditions between *j3* and *j4.* These results suggest that P3a amplitudes form an “inverted-U” with differential entropy of the stimulus, peaking at moderate entropy values while remaining low for both low and high entropy values.

P3 latency was not significantly related to sound harmonicity, both in the ANOVA analysis ($F(7, 238) = 1.22$, $p_{Greenhouse–Geisser} = 0.299$, see also ST2) and in linear and polynomial mixed models (all p-values > .05). Model comparison results and parameters of the fitted models are reported in ST4 and ST5.

Behavioral results

In the behavioral experiment, participants were required to perform a two-alternative forced choice pitch discrimination task. For a sequence of two sounds, the participant was to determine if the second sound is higher or lower in pitch than the first sound. We employed an adaptive staircase procedure that modulated the amount of inharmonicity applied to the sounds. A psychometric curve fitted on the behavioral data suggests that the (75%) detection threshold for the task was 5.09 (95% C.I. 4.98 - 5.39), indicating that pitch discrimination for inharmonic sounds deteriorated around *j5*. Notably, the upper asymptote was estimated at 75% correct responses (lambda = .243, 95% C.I. .211 - .244) indicating that the task proved difficult even for harmonic or slightly inharmonic sounds.

Discussion

The results of this study show that the relationship between the level of inharmonicity and MMN amplitude can be modelled best with a sigmoid function. This is additionally supported by ANOVA and post-hoc contrast results, showing that MMN amplitude is similar for harmonic and low-inharmonic sounds. At the same time, MMN amplitudes are significantly smaller for highly inharmonic sounds. The sigmoid model parameters and post-hoc comparisons place the threshold above which MMN amplitude suddenly decreases to be between jitters *j4* and *j5*. Converging evidence from the behavioral experiment suggests that this threshold also signifies a level of inharmonicity above which pitch discrimination is severely impaired. Overall, the presented results support the hypothesis that inharmonicity does not exert influence on auditory deviance detection until a threshold where it begins to impair F0 extraction. At the same time, a

lack of gradual decrease in MMN amplitudes is in conflict with theories that suggest that increased information entropy leads to down-weighting of prediction errors by precision.

It is possible that precision-weighting processes are not reflected in the MMN amplitudes or that precision-induced changes are so minute that they are not observable in noisy EEG data. While this is certainly a possibility, recent simulation work suggests that precision weighting should not only be apparent in amplitudes, but also in the latencies of the putative event-related potentials (Bowman et al., 2023). We do not find this to be the case in our results, as latencies of both MMN and P3a were not modulated by the level of inharmonicity. However, one needs to remember that different models of how precision-weighting might be implemented in the brain exist in the literature (i.e. gain modulation, variability modulation, synaptic modulation, (Hertäg et al., 2025) see also (Aizenbud et al., 2025) for review).

While our results suggest that harmonicity is not an index for precision-weighting, they do not contradict the models of predictive processing that implement precision-weighting. Rather, they suggest a process where pitch information is extracted from harmonic to mildly inharmonic sounds, but the auditory system struggles to perform this task for higher inharmonicity values. Further, the confidence (precision) of any pitch related predictions does not seem to influence the amplitude or latency of ERP components. This possibly results from the fact that F0 extraction / prediction might be performed earlier in the auditory processing hierarchy.

We also observed a slight peak in the P3a responses around the inharmonicity threshold which can reflect the higher computational demand associated with detecting pitch deviants at the threshold level. Interestingly, this peak occurred for slightly lower inharmonicity levels than the MMN sigmoid threshold, indicating that this computational demand is the highest for near-threshold inharmonicity levels.

It is important to remember that the results of the present study were obtained from a passive paradigm. This likely contributed to the overall small ERP amplitudes (REF). With an active task (i.e., respond to pitch deviants), a P3b response might have also been elicited, exerting influence on the P3a. Furthermore, the computations performed by the auditory system might be different in an active task in comparison to a passive one, which could lead to different patterns of mismatch responses (particularly in the P300 complex). Different computations might also underpin different perceptual strategies for detecting deviants, with substantial inter-individual differences. Future research could address these issues by utilizing active paradigms on larger sample sizes, perhaps facilitated by gamification to counter participant fatigue and support motivation.

We acknowledge that the behavioral results presented in this study are limited. The aim of the behavioral experiment was to gain insight into the detectability of pitch deviants in order to better interpret the EEG results. As such, the behavioral experiment was not designed to provide thorough psychometric testing of pitch discrimination under inharmonicity. Rather, our aim was to verify if the pitch deviants presented in the passive EEG experiment could be detected if participant's attention was afforded. Specifically, we have made a decision to retain a

constant 50 Hz difference between sounds for each pair in the range of 200 - 500 Hz. We recognize that this decision resulted in different perceived pitch intervals for different starting frequencies, with a bias towards smaller intervals for higher pitches. However, exactly the same conditions were present in the roving paradigm and we wanted to make sure that the behavioral results are, first and foremost, interpretable in the context of the EEG results. A full examination of the effects of inharmonicity on pitch discrimination would involve testing different intervals independently of different inharmonicity levels and is outside the scope of this study.

Methods

Participants

37 people signed up for the study. Inclusion criteria were: no reported neurological or psychiatric illness, age between 18 and 45 years old, normal hearing, and no use of medication that affects the central nervous system 24 hours prior to testing (e.g., opioids, pain medications). The participants were subjected to a tonal audiometry test before the experiment to screen for hearing disorders. Two of the participants did not meet the inclusion criteria, with hearing thresholds below the -20 dB HL (hearing level) benchmark. The remaining 35 did not have problems with hearing and took part in the study. Two participants had to perform the experiment in two sessions, due to tension headaches from the EEG and headphones placed on their heads. The participants were aged between 18 and 37 years (M = 23.7, SD = 3.17). 27 (77%) of them were women, 8 (23%) were men. Participants received monetary compensation for their involvement in the study.

Before data collection, we performed a power analysis to guide sample size decisions. We have simulated MMN peak amplitude data assuming a worst-case scenario where all conditions did not differ except one that was, on average, 1 μV larger than the rest (a difference of this magnitude was observed between conditions in (Basiński et al., 2025)). We ran this simulation 10 000 times for each sample size in the range between 20 and 40. For each simulated dataset, we fit a mixed-models ANOVA with condition as a fixed effect and participant as a random effect. The results of these simulations revealed that in order to achieve statistical power of 0.8 or more, at least N = 34 participants were required (ST 6).

Procedure

The protocol for this study was approved by the Scientific Research Bioethics Committee of the Medical University of Gdańsk (reference number KB/489/2024) and followed the Declaration of Helsinki.

Each participant took part in the following procedures: (1) tonal audiometry - in order to check that the participant has healthy hearing, (2) the behavioral experiment and (3) the EEG experiment. Experiments were performed on a computer running Psychopy v.2023.2.3 (for behavioral) or v.2024.2.1 (for EEG). The sound stimuli were presented binaurally with Beyerdynamic DT 770 Pro headphones and an RME UCX II audio interface.

In the behavioral experiment, participants performed a two-alternative forced choice pitch discrimination task. They were listening to pairs of sounds (IOI = 500 ms) and were asked to determine whether the second sound was higher or lower in pitch than the first sound. The level of inharmonicity was manipulated with a 3-up 1-down staircase procedure. Pairs of sounds were presented with an inter-onset interval of 500 ms. The sounds' F0s were randomly chosen from a range of 200 Hz - 500 Hz in increments of 10 Hz. The difference in frequency between sounds was always set at 50 Hz in order to ensure correspondence with the roving paradigm (see below). Each staircase block ended after 6 reversals and each participant completed 15 blocks. Before the main experiment, a familiarization trial was run with a similar task and using harmonic sounds. The participant proceeded to the main task after two consecutive correct responses in the familiarization trial. Participants responded by pressing a button on the keyboard.

Scalp EEG potentials were collected with a 64-channel active system (BioSemi B.V.) with two additional EOG electrodes. The EEG experiment consisted of 16 blocks, two per inharmonicity condition, each containing roving sequences of 600 sounds played with an ISI of 600 ms. The first tone of a new stimulus train served as the deviant stimulus and a source of a mismatch response in the context of the prior stimulus train. After a few repetitions, the deviant tone establishes itself as the new standard until the next F0 change happens. The number of stimuli in a given train varied pseudo-randomly from 4 to 7. Each block lasted about 6 minutes. The F0s were chosen from the set of frequencies from 200 Hz to 500 Hz, spaced 50 Hz apart. The participants were asked to refrain from movement during sound stimulation but could take self-paced breaks between the blocks. During EEG recording, the participants watched a movie of their choice, without sound. The entire procedure, including behavioral testing, lasted for about 2 hours.

Stimuli

In both experiments, stimuli consisted of synthetic complex tones generated with custom Python code, based on similar procedures used in previous studies (Basiński et al., 2025; McPherson & McDermott, 2018, 2020). Each tone was synthesized by adding sine waves at frequencies of F0 and its harmonic frequencies up to the Nyquist limit (24 kHz), in random phase. For inharmonic sounds, random jitter *J* was applied to each frequency above the F_0. The jittered frequencies F_n' were calculated using the following formula:

$$F_n' = F_n(1 + J),\ J \sim U(-j_c, j_c)$$

where F_n is the frequency of the n-th harmonic and J is a continuous uniformly distributed random variable ranging from $-j_c$ to j_c, with j_c being the condition-specific jitter rate. Rejection sampling was used in cases where two neighbouring frequencies differed by less than 30 Hz in order to avoid beating artifacts. For every sound, a fourth-order Butterworth high-pass filter was applied at the cutoff frequency of 3500 Hz, corresponding to the 10th harmonic of the average F0 used throughout both experiments (350Hz). This filter provided attenuation of the F0 in order to minimize the variation at the spectral edge of the fundamental (see (McPherson & McDermott, 2018) for a discussion of the rationale behind this approach). Sounds had a length of 70 ms and were amplitude modulated with 5 ms onset and offset ramps (half-Hanning windows). All sounds were pre-rendered as 16-bit wav files with a sampling rate of 48 kHz and loudness-normalized to -12 LUFS using *pyloudnorm*. For the behavioral experiment, 100 sounds with different jitter patterns were created for each possible F0 and each jitter rate, for a total of 35,000 sounds. For the EEG experiment, 1000 sounds with different jitter patterns were created for each possible F0 and jitter pattern, for a total of 56,000 sounds.

The blocks consisted of sounds with exponentially growing jitters ranging from 0 to 0.108. The exact jitter values were as follows: 0, 0.005, 0.008, 0.014, 0.023, 0.039, 0.065, 0.108. Each set was presented twice.

Data processing and statistical analysis

The ERP data were collected using a 64-channel BIOSEMI system at a sampling rate of 2048 Hz. All data was decimated by a factor of two before preprocessing and subsequent analyses were performed with a sample rate of 1024 Hz. The oculomotor activity was recorded using two additional EOG electrodes. Raw EEG data underwent high-pass (0.2 Hz), low-pass (30 Hz), and notch (50 Hz) filtering to remove the AC component. The data were then segmented into epochs ranging from -100 ms to 500 ms and subsequently processed using the *autoreject* algorithm v0.4.2 (Jas et al., 2017) for automatic detection and rejection of faulty channels and epochs. This machine learning algorithm estimates optimal peak-to-peak signal thresholds for each epoch and each EEG sensor. The thresholding leads to epoch-wise interpolation of bad channels or (in case of very noisy signals) epoch rejection. Next, in order to remove oculomotor artifacts, an independent component analysis (ICA) was performed. ICA components were identified for removal through an automated procedure based on the EOG signals; however, all selections were later manually verified. The ICA solution was applied to the data before automatic artifact rejection. Then, the autoreject algorithm was run once again, as recommended by Jas et al. (2017). In the end, 13 400 out of 321 389 of epochs had to be rejected (4.2%). The remaining epochs were re-referenced to the mastoids and baseline-corrected using a 100 ms pre-stimulus interval.

A sound was classified as a deviant if it was the first occurrence of a given F0 following a change in the roving paradigm. Five initial sounds in the sequence and every two sounds following a deviant were rejected from the analysis, with the remaining sounds classified as standards. ERPs were computed by averaging responses to standard and to deviant sounds

within participants and conditions. Difference waves were derived by subtracting the ERPs of standard sounds from those of deviant sounds for each condition and participant. Peak measures were calculated on averaged EEG signal from fronto-central electrodes Fz, F1, F2, F3, F4, AF3, AF4, AFz, FC1, FC2 and FCz. MMN latencies were calculated participant- and condition-wise by detecting the latency of the most negative peak (local minimum) in the window from 70 ms to 250 ms post-stimulus. Similarly, P3a latencies were calculated for a positive peak in the window between 150 ms and 350 ms. MMN and P3a amplitudes were calculated as mean amplitudes of the signal between +/- 25 ms around the participant- and condition-wise peak.

We tested whether the MMN and P3 amplitudes differ significantly from zero using one-sample, two-tailed t-tests. *P*-values correction based on the Benjamini-Hochberg procedure was applied to keep the false discovery rate at the 0.05 significance level, separately for the MMN and P3 amplitudes comparisons. To test whether harmonicity affects the amplitude and latency of mismatch responses, we performed repeated measures ANOVA. Greenhouse–Geisser correction was applied whenever Mauchly's Test for Sphericity yielded significant results. If ANOVA results showed a significant effect of harmonicity, we performed pairwise post-hoc comparisons using estimated marginal means. P-values were corrected for multiple comparisons using the Tukey method.

To model how increasing inharmonicity affects mismatch responses, we fitted mixed effects models, with differential entropy of the jitter rates distribution, specific to the experimental condition, as the independent variable. Polynomial models up to the fourth degree with random intercepts were fitted for both MMN and P3 amplitude and each model was compared to the lower degree model with a likelihood ratio test. To test the hypothesis that MMN amplitude remains at a similar level for low inharmonicity conditions and then rapidly approaches zero we fitted a mixed effects sigmoid model with random asymptote *L*:

$$MMN_{amplitude} = \frac{L}{1+e^{k(I-I_0)}},$$

where I is the differential entropy of the jitter rates distribution (calculated using log_{10}), I_0 is the I value at the inflection point, and *k* determines the slope at the inflection point.

All EEG signal processing was performed in Python v3.11, using the packages: MNE v1.5 (Gramfort, 2013; Larson et al., 2024), Numpy v1.24 (Harris et al., 2020), SciPy v1.11 (Virtanen et al., 2020) and Pandas v2.1 (McKinney, 2010). Data visualization was carried out using Matplotlib v3.7. Modelling and statistical tests were performed in R version 4.3.3 (R Core Team, 2024), with *lme4* package, version 1.1.35.5 (Bates et al., 2015) used for mixed model fitting and *ez* package, version 4.4.0 (Lawrence, 2016) used for ANOVA.

Acknowledgements

This work was supported by a grant from the National Science Centre, Poland (Sonata, no. 2022/47/D/HS6/03323).

Author contributions

A.B.: formal analysis, investigation, writing - review & editing.
B.W.: software, investigation, writing - review & editing.
T.D.: investigation, writing - review & editing.
M.B. - formal analysis, visualisation, writing - original draft, writing - review & editing.
KB: conceptualization, methodology, software, data analysis, data curation, funding acquisition, writing - original draft, review & editing, visualization, supervision.

Data availability statement

The raw data for this experiment is available at xxxx while code that was used to perform the analyses is available at xxxx.

Competing interests statement

The authors declare no competing interests.

References

Aizenbud, I., Audette, N., Auksztulewicz, R., Basiński, K., Bastos, A. M., Berry, M., Canales-Johnson, A., Choi, H., Clopath, C., Cohen, U., Costa, R. P., Filippo, R. D., Doronin, R., Errington, S. P., Gavornik, J. P., Gillon, C. J., Granier, A., Hamm, J. P., Hertäg, L., … Xiong, Y. S. (2025). Neural mechanisms of predictive processing: A collaborative community experiment through the OpenScope program (No. arXiv:2504.09614). arXiv. https://doi.org/10.48550/arXiv.2504.09614

https://doi.org/10.1016/j.pneurobio.2024.102696

Aizenbud, I., Audette, N., Auksztulewicz, R., Basiński, K., Bastos, A. M., Berry, M., Canales-Johnson, A., Choi, H., Clopath, C., Cohen, U., Costa, R. P., Filippo, R. D., Doronin, R., Errington, S. P., Gavornik, J. P., Gillon, C. J., Granier, A., Hamm, J. P.,

Hertäg, L., … Xiong, Y. S. (2025). *Neural mechanisms of predictive processing: A collaborative community experiment through the OpenScope program* (No. arXiv:2504.09614). arXiv. https://doi.org/10.48550/arXiv.2504.09614

Basiński, K., Celma-Miralles, A., Quiroga-Martinez, D. R., & Vuust, P. (2025). *Inharmonicity enhances brain signals of attentional capture and auditory stream segregation* (p. 2025.04.17.649377). bioRxiv. https://doi.org/10.1101/2025.04.17.649377

Basiński, K., Quiroga-Martinez, D. R., & Vuust, P. (2023). Temporal hierarchies in the predictive processing of melody − From pure tones to songs. *Neuroscience & Biobehavioral Reviews*, *145*, 105007. https://doi.org/10.1016/j.neubiorev.2022.105007

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bowman, H., Collins, D. J., Nayak, A. K., & Cruse, D. (2023). Is predictive coding falsifiable? *Neuroscience & Biobehavioral Reviews*, *154*, 105404. https://doi.org/10.1016/j.neubiorev.2023.105404

Carbajal, G. V., & Malmierca, M. S. (2018). The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends in Hearing*, *22*, 233121651878482. https://doi.org/10.1177/2331216518784822

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory Prerequisites of Mismatch Negativity in the Auditory Event-Related Potential (ERP). *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *19*, 909–921. https://doi.org/10.1037/0278-7393.19.4.909

de Cheveigné, A. (2021). Harmonic Cancellation—A Fundamental of Auditory Scene Analysis. *Trends in Hearing*, *25*, 23312165211041422.

https://doi.org/10.1177/23312165211041422

Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Garrido, M. I., Friston, K. J., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Kilner, J. M. (2008). The functional anatomy of the MMN: A DCM study of the roving paradigm. *NeuroImage*, *42*(2), 936–944. https://doi.org/10.1016/j.neuroimage.2008.05.018

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, 11.

Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, *389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Hertäg, L., Wilmes, K. A., & Clopath, C. (2025). Uncertainty estimation with prediction-error circuits. *Nature Communications*, *16*, 3036. https://doi.org/10.1038/s41467-025-58311-6

Kanai, R., Komura, Y., Shipp, S., & Friston, K. (2015). Cerebral hierarchies: Predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1668), 20140169. https://doi.org/10.1098/rstb.2014.0169

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron*, *100*(2), 424–435. https://doi.org/10.1016/j.neuron.2018.10.003

Lawrence, M. A. (2016). *ez: Easy Analysis and Visualization of Factorial Experiments* (Version 4.4-0) [Computer software]. https://cran.r-project.org/web/packages/ez/index.html

McPherson, M. J., Grace, R. C., & McDermott, J. H. (2022). Harmonicity aids hearing in noise. *Attention, Perception, & Psychophysics*, *84*(3), 1016–1042. https://doi.org/10.3758/s13414-021-02376-0

McPherson, M. J., & McDermott, J. H. (2018). Diversity in pitch perception revealed by task dependence. *Nature Human Behaviour*, *2*(1), 52–66.

https://doi.org/10.1038/s41562-017-0261-8

McPherson, M. J., & McDermott, J. H. (2020). Time-dependent discrimination advantages for harmonic sounds suggest efficient coding for memory. *Proceedings of the National Academy of Sciences*, *117*(50), 32169–32180. https://doi.org/10.1073/pnas.2008956117

Milne, A. E., Zhao, S., Tampakaki, C., Bury, G., & Chait, M. (2021). Sustained Pupil Responses Are Modulated by Predictability of Auditory Sequences. *The Journal of Neuroscience*, *41*(28), 6116–6127. https://doi.org/10.1523/JNEUROSCI.2879-20.2021

Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, *66*(3), 241–251. https://doi.org/10.1007/BF00198477

Palmer, C. E., Auksztulewicz, R., Ondobaka, S., & Kilner, J. M. (2019). Sensorimotor beta power reflects the precision-weighting afforded to sensory prediction errors. *NeuroImage*, *200*, 59–71. https://doi.org/10.1016/j.neuroimage.2019.06.034

Polich, J., & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology*, *60*(2), 172–185. https://doi.org/10.1016/j.ijpsycho.2005.12.012

Popham, S., Boebinger, D., Ellis, D. P. W., Kawahara, H., & McDermott, J. H. (2018). Inharmonic speech reveals the role of harmonicity in the cocktail party problem. *Nature Communications*, *9*(1), 2122. https://doi.org/10.1038/s41467-018-04551-8

Quiroga-Martinez, D. R., Basiński, K., Nasielski, J., Tillmann, B., Brattico, E., Cholvy, F., Fornoni, L., Vuust, P., & Caclin, A. (2022). Enhanced mismatch negativity in harmonic compared to inharmonic sounds. *European Journal of Neuroscience*, *n/a*(n/a). https://doi.org/10.1111/ejn.15769

R Core Team. (2024). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org

Rao, R. P. N. (2024). A sensory–motor theory of the neocortex. *Nature Neuroscience*, *27*(7), 1221–1235. https://doi.org/10.1038/s41593-024-01673-9

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, *83*(2), 132–143. https://doi.org/10.1016/j.ijpsycho.2011.10.001

Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology*, *31*(17), R1026–R1032. https://doi.org/10.1016/j.cub.2021.07.044

Zhao, S., Skerritt-Davis, B., Elhilali, M., Dick, F., & Chait, M. (2025). Sustained EEG responses to rapidly unfolding stochastic sounds reflect Bayesian inferred reliability tracking. *Progress in Neurobiology*, *244*, 102696. https://doi.org/10.1016/j.pneurobio.2024.102696

Supplementary materials

Supplementary Table 1. Mean MMN and P3 amplitudes, tested against 0.

Condition	*M[95%CI]*	*t*	*df*	*p*	p_{adj}
MMN amplitude					
j0	-1.17[-1.61, -0.74]	-5.45	34	<.001	<.001
j1	-1.49[-1.86, -1.11]	-8.08	34	<.001	<.001
j2	-1.13[-1.51, -0.76]	-6.18	34	<.001	<.001
j3	-1.22[-1.66, -0.78]	-5.61	34	<.001	<.001
j4	-0.91[-1.28, -0.54]	-4.97	34	<.001	<.001
j5	-0.37[-0.62, -0.12]	-2.97	34	0.005	0.007
j6	-0.39[-0.73, -0.06]	-2.39	34	0.023	0.026
j7	-0.32[-0.64, -0.01]	-2.08	34	0.045	0.045
P3 amplitude					
j0	0.93[0.54, 1.32]	4.83	34	<.001	<.001
j1	0.51[0.12, 0.89]	2.69	34	0.011	0.011
j2	1.01[0.6, 1.42]	5.00	34	<.001	<.001
j3	0.91[0.55, 1.27]	5.14	34	<.001	<.001
j4	0.99[0.57, 1.41]	4.75	34	<.001	<.001
j5	0.79[0.51, 1.07]	5.71	34	<.001	<.001
j6	0.54[0.19, 0.88]	3.17	34	0.003	0.004
j7	0.58[0.24, 0.92]	3.48	34	0.001	0.002

Note: One sample two-sided t tests were used. Results were considered significant if $p_{adj} < 0.05$; p_{adj} was calculated using Benjamini & Hochberg correction; j0 is the control, harmonic condition and j1-7 indicate subsequent inharmonic conditions.

Supplementary Table 2. ANOVA results for the effect of inharmonicity on MMN and P3 amplitudes and latencies.

Effect	DFn	DFd	SSn	SSd	F	p	ges	W	p_{GG}
MMN amplitude									
(Intercept)	1	34	214.39	121.52	59.99	<.001	0.41		
jitter	7	238	50.60	188.55	9.12	<.001	0.14	0.17*	<.001
P3 amplitude									
(Intercept)	1	34	170.46	106.25	54.54	<.001	0.35		
jitter	7	238	10.81	206.97	1.78	0.093	0.0	0.37	

MMN latency									
(Intercept)	1	34	6.85	0.05	4363.87	<.001	0.94		
jitter	7	238	0.01	0.35	0.82	0.57	0.02	0.2*	0.538
P3 latency									
(Intercept)	1	34	18.40	0.09	7223.28	<.001	0.98		
jitter	7	238	0.01	0.36	1.22	0.29	0.03	0.23*	0.299

Supplementary Table 3. Contrasts of MMN amplitudes in inharmonic conditions (j1-j7) with harmonic condition j0.

Condition	*M[95%CI]*	*t*	*df*	*p*	p_{adj}
j1	-0.31[-0.62, 0]	-2.04	34	0.049	0.086
j2	0.04[-0.31, 0.39]	0.24	34	0.812	0.812
j3	-0.04[-0.4, 0.31]	-0.25	34	0.803	0.812
j4	0.27[-0.13, 0.66]	1.37	34	0.179	0.251
j5	0.81[0.37, 1.25]	3.70	34	0.001	0.005
j6	0.78[0.28, 1.28]	3.19	34	0.003	0.007
j7	0.85[0.32, 1.39]	3.24	34	0.003	0.007

Note: One sample two-sided t tests were used. Results were considered significant if $p_{adj} < 0.05$; p_{adj} was calculated using Benjamini & Hochberg correction; j0 is the control, harmonic condition and j1-7 indicate subsequent inharmonic conditions.

Supplementary Table [?]. Pairwise contrasts of mismatch negativity amplitude across jitter conditions.

Jitter pair	Estimate	SE	df	T. ratio	p
j0 - j1	0.3108	0.213	238	1.461	0.827
j0 - j2	-0.0416	0.213	238	-0.196	1.000
j0 - j3	0.0435	0.213	238	0.204	1.000
j0 - j4	-0.2676	0.213	238	-1.258	0.913
j0 - j5	-0.8093	0.213	238	-3.804	0.004
j0 - j6	-0.7823	0.213	238	-3.677	0.007
j0 - j7	-0.8521	0.213	238	-4.005	0.002

j1 - j2	-0.3524	0.213	238	-1.656	0.715
j1 - j3	-0.2673	0.213	238	-1.256	0.914
j1 - j4	-0.5784	0.213	238	-2.718	0.122
j1 - j5	-1.1201	0.213	238	-5.264	<.001
j1 - j6	-1.0931	0.213	238	-5.137	<.001
j1 - j7	-1.1629	0.213	238	-5.465	<.001
j2 - j3	0.0851	0.213	238	0.400	1.000
j2 - j4	-0.2260	0.213	238	-1.062	0.964
j2 - j5	-0.7677	0.213	238	-3.608	0.009
j2 - j6	-0.7407	0.213	238	-3.481	0.014
j2 - j7	-0.8105	0.213	238	-3.809	0.004
j3 - j4	-0.3111	0.213	238	-1.462	0.827
j3 - j5	-0.8528	0.213	238	-4.008	0.002
j3 - j6	-0.8258	0.213	238	-3.881	0.003
j3 - j7	-0.8956	0.213	238	-4.209	0.001
j4 - j5	-0.5417	0.213	238	-2.546	0.182
j4 - j6	-0.5147	0.213	238	-2.419	0.237
j4 - j7	-0.5845	0.213	238	-2.747	0.114
j5 - j6	0.0270	0.213	238	0.127	1.000
j5 - j7	-0.0428	0.213	238	-0.201	1.000
j6 - j7	-0.0698	0.213	238	-0.328	1.000

Note: Pairwise comparisons were conducted using estimated marginal means (EMMs) to evaluate differences in mismatch negativity (MMN) amplitude between jitter conditions (j0–j7). The table reports the contrast estimates (in μV) and p-values. p-values were adjusted for multiple comparisons using the Tukey method.

Supplementary Table 4. Comparison of models of MMN and P3 amplitude as a function of jitter differential entropy.

Model degree	df	AIC	BIC	logLik	Test	L.Ratio	p
MMN amplitude - linear and polynomial models							
0 (Intercept only)	3	736.13	746.63	-365.07		-	-
1 (linear)	4	691.23	705.24	-341.62	0 vs 1	46.90	<.001
2 (quadratic)	5	693.01	710.51	-341.50	1 vs 2	0.23	0.635
3 (cubic)	6	693.66	714.67	-340.83	2 vs 3	1.35	0.246
4 (quartic)	7	694.77	719.28	-340.38	3 vs 4	0.89	0.345

MMN amplitude - sigmoid model							
0 (Intercept only)	3	736.13	746.63	-365.07		-	-
Sigmoid model	5	664.65	682.15	-327.32	1 vs 2	75.48	<.001
P3 amplitude - linear and polynomial models							
0 (Intercept only)	3	716.43	726.93	-355.22		-	-
1 (linear)	4	717.46	731.46	-354.73	0 vs 1	0.97	0.324
2 (quadratic)	5	713.50	731.01	-351.75	1 vs 2	5.95	0.015
3 (cubic)	6	712.62	733.63	-350.31	2 vs 3	2.88	0.090
4 (quartic)	7	714.62	739.13	-350.31	3 vs 4	0.00	0.966

Supplementary Table 5. Fitted models fixed effects.

	Value	**Std.Error**	**DF**
MMN amplitude - polynomial model			
Intercept	-0.83	0.1	206
Differential entropy (I)	6.51	0.9	206
Differential entropy (I) 2nd power	-0.43	0.9	206
Differential entropy (I) 3rd power	-1.04	0.9	206
Differential entropy (I) 4th power	-0.84	0.9	206
MMN amplitude - sigmoid model			
Asymptote (L)	-1.31	0.18	208
Inflection point (I_0)	-1.18	0.05	208
Slope at the inflection point (k)	6.24	1.73	208
P3 amplitude - polynomial model			
Intercept	0.76	0.10	206
Differential entropy (I)	-0.95	0.95	206
Differential entropy (I) 2nd power	-2.33	0.95	206
Differential entropy (I) 3rd power	1.60	0.95	206
Differential entropy (I) 4th power	0.04	0.95	206

Note: polynomial models utilized orthogonal polynomials; thus, for polynomial models of degree lower than 4, estimates and standard errors were the same as in the 4th degree model.