



OPEN

Evoked responses to note onsets and phrase boundaries in Mozart's K448

Yijing Feng^{1,5}, Robert J. Quon^{2,3,5}, Barbara C. Jobst^{2,3,5} & Michael A. Casey^{1,4,5}✉

Understanding the neural correlates of perception of hierarchical structure in music presents a direct window into auditory organization. To examine the hypothesis that high-level and low-level structures—i.e. phrases and notes—elicit different neural responses, we collected intracranial electroencephalography (iEEG) data from eight subjects during exposure to Mozart's K448 and directly compared Event-related potentials (ERPs) due to note onsets and those elicited by phrase boundaries. Cluster-level permutation tests revealed that note-onset-related ERPs and phrase-boundary-related ERPs were significantly different at –150, 200, and 450 ms relative to note onset and phrase markers. We also observed increased activity in frontal brain regions when processing phrase boundaries. We relate these observations to (1) a process which syntactically binds notes together hierarchically to form larger phrases; (2) positive emotions induced by successful prediction of forthcoming phrase boundaries and violations of melodic expectations at phrase boundaries.

Musical information is organized hierarchically. The processing of individual musical elements such as phrase boundaries and note onsets, is associated with distinct brain regions and neural responses. Understanding the neural correlates of perception of hierarchical structure in music presents a direct window into auditory organization.

The music-theoretic concept of musical structure describes listeners' segmentation of auditory information into nested hierarchical units of various sizes¹. Previous work such as Lerdahl and Jackendoff's *A Generative Theory of Tonal Music*², which was influenced by Berstein's *The Unanswered Question*³, attempted to model music understanding with the aid of generative linguistics. In principle, the organization in music is similar to human language, where speech is nested recursively into units such as phonemes and words, and extended to phrases and sentences. Ding et al.⁴ have shown a hierarchy of neural processing timescales underlies grammar-based internal construction of hierarchical linguistic structure. Prystauka et al.⁵ reviewed recent studies and summarized the theories linking the oscillatory markers to the processing of hierarchical structure in languages, such as linking beta oscillation to syntactic structure building and linking gamma oscillation to semantic structure building^{6,7}. Correspondingly, music consists of notes, chords, themes, and higher-level functional units such as phrases and sections⁸, which occur at quasi-periodic intervals and are marked by changes in melodic theme, harmony, rhythm, and key^{9,10}. These higher-level compositional elements underlie audience engagement with the music and are experienced as anticipation of upcoming events. Thus, phrase-level components are regarded as primary functional units in the cognitive processing of music.

To better understand the cognitive processing of complex auditory information, previous studies have investigated neural responses to important structural elements in music by examining event-related potentials (ERPs). Several ERP components that are linked to syntactic violations in language processing have also been observed in music perception. For example, the N400 component is associated with words that are semantically anomalous given the preceding context¹¹, and it was also discovered to be elicited in the processing of out-of-key or unexpected notes in familiar melodies^{12,13}. The P600 component, which is sensitive to the non-preferred continuation of a sentence¹⁴, can also be elicited by incongruous elements in musical sequences^{15,16}. In addition, the closure positive shift (CPS), an electrical phenomenon that can be detected at the close of a phrase, has been reported to mark prosodic phrase boundaries in both speech¹⁷ and music¹⁸. These findings contribute to the understanding of the perception of individual higher-order structural elements in music.

¹Department of Computer Science, Dartmouth College, Hanover, NH 03755, USA. ²Geisel School of Medicine, Dartmouth College, Hanover, NH 03755, USA. ³Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, USA. ⁴Department of Music, Dartmouth College, Hanover, NH 03755, USA. ⁵These authors contributed equally: Yijing Feng, Robert J. Quon, Barbara C. Jobst and Michael A. Casey. ✉email: michael.a.casey@dartmouth.edu

Subregion	Number of temporal clusters	Mean of <i>p</i> -value	SD of <i>p</i> -value
Superior temporal cortex	21	0.010	0.012
Rostral anterior cingulate cortex	10	0.015	0.017
Rostral middle frontal cortex	14	0.024	0.015
Medial orbitofrontal cortex	11	0.014	0.016

Table 1. Statistics of significant temporal clusters in the comparison between iEEG sampled around note onsets and reference windows within each subregion in all twelve sessions.

Subregion	Number of temporal clusters	Mean of <i>p</i> -value	SD of <i>p</i> -value
Middle temporal cortex	60	0.012	0.014
Superior temporal cortex	57	0.012	0.013
Rostral anterior cingulate cortex	63	0.011	0.013
Rostral middle frontal cortex	77	0.013	0.014
Medial orbitofrontal cortex	49	0.013	0.014
Superior frontal cortex	62	0.015	0.014
Insular cortex	61	0.012	0.014
Caudal middle frontal cortex	69	0.013	0.014

Table 2. Statistics of significant temporal clusters in the comparison between iEEG sampled around phrase boundaries and reference windows within each subregion in all twelve sessions.

However, it remains unclear how the human brain processes and integrates auditory information at different hierarchical levels with naturalistic music stimuli. Most previous studies extracted musical phrases from simple melodies or manipulated phrase boundaries by note filling—a commonly used technique to generate unphrased control stimuli by filling pauses with musically plausible notes, which do not allow for the investigation of the neural processing of phrase boundaries in naturalistic music perception. Other studies attempted to explore hierarchy in music perception but failed to reveal the neural correlates of higher-order structural elements due to the lack of score-based segmentation of musical stimuli. These studies relied on neural responses to the non-cognitive units marked by pauses or bars¹⁹, which limited their findings to the lower-level perception of music.

To address the gap in understanding the neural correlates of different hierarchical levels of music perception, we analyzed brain responses to naturalistic music with note-onset and phrase-boundary-related ERPs using a cluster-based permutation test, and localized brain structures activated by these different stimulus elements. The current study extends previous work in two ways: (1) it directly compared the neural responses to musical components at different levels, which helps reveal the hierarchical structure in auditory cognition, and (2) it generalized Knösche's result¹⁸ to naturalistic music perception by using naturalistic, i.e. unmodified, musical-phrase stimuli. We hypothesized that low-level and high-level musical structures would elicit distinct neural responses, and that the processing of low-level structures would be associated with lateral temporal brain regions and high-level structures would involve increased activity in frontal brain regions. As such, our study serves as a foundation for understanding brain responses to the hierarchical structure in music perception.

Results

Note-onset-related ERPs and phrase-boundary-related ERPs. A total of twelve sessions of Intracranial Stereo-EEG data were collected from eight subjects with refractory epilepsy undergoing intracranial EEG monitoring for the clinical treatment during exposure to the first 90 s of Mozart's K448.

To verify the hypothesis that both note onsets and phrase boundaries elicit evoked responses, we computed the ERP waveforms (Figs. 1, 2) by averaging intracranial electroencephalography (iEEG) windows sampled near stimulus markers (phrase boundaries and note onsets) across all twelve sessions, and performed a cluster-based permutation test to determine whether the iEEG windows sampled near the stimulus markers were significantly different from reference windows randomly sampled between stimulus markers. Temporal clusters in which a significant difference ($p < 0.05$) was observed between the two conditions were reported. The *p*-value statistics of significant temporal clusters are provided in Tables 1 and 2. Figure 3 shows the result of the statistical analysis in each subregion within a single session. We shaded the regions in Figs. 1 and 2 to represent the intersection of significant temporal clusters across all twelve sessions in each subregion.

We then integrated the results for frontal brain regions and lateral temporal brain regions over all twelve sessions as shown in Figs. 4 and 5. For note-onset-related ERPs, a majority of sessions had significant temporal clusters at −150 and 200 ms around the note onset stimulus markers in both frontal and temporal brain regions. Notably, ten sessions contained significant clusters at around 200 ms. In the analysis of phrase-boundary-related ERPs, all twelve sessions having overlapping clusters at −150, 0, 200, and 400 ms around the phrase boundaries in both frontal and temporal brain regions. Figures 6 and 7 illustrate the cortical distribution of the statistical

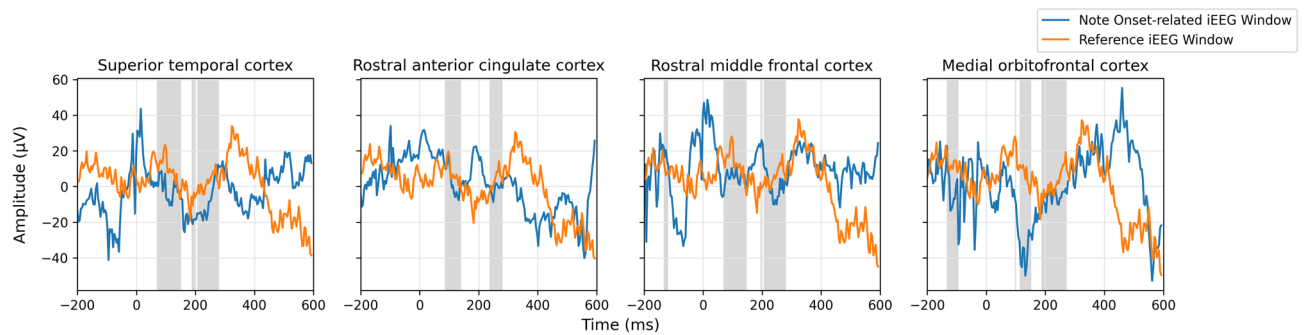


Figure 1. Note-Onset-related ERPs observed at these subregions with significance in at least six sessions are shaded in light gray.

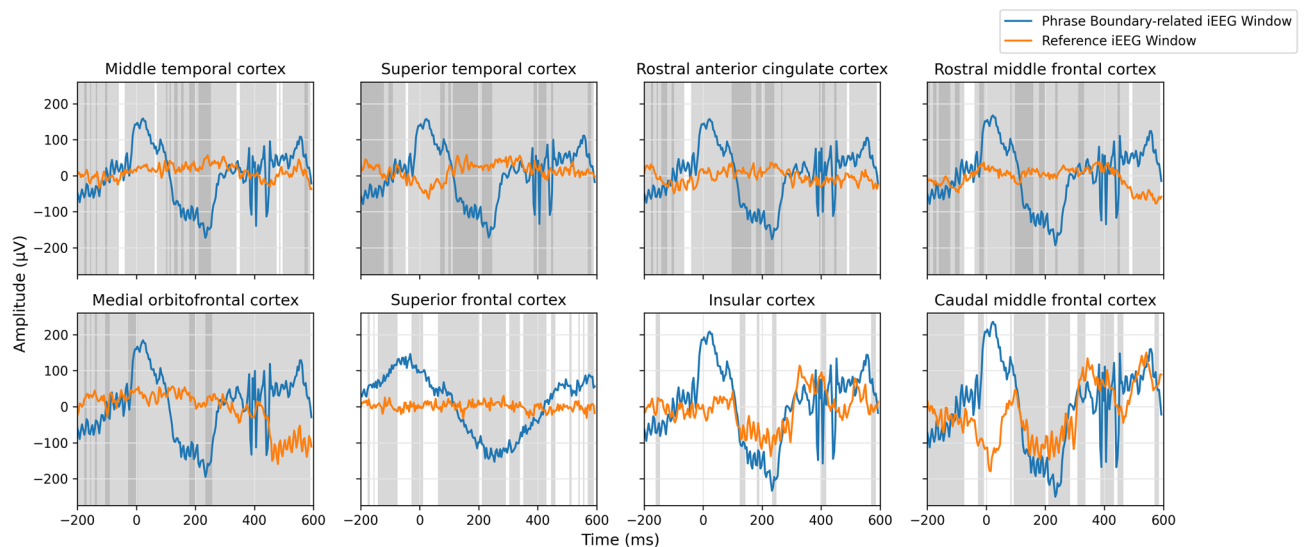


Figure 2. Phrase-boundary-related ERPs observed at these subregions with significance in at least six sessions are shaded in light gray and ten sessions in dark gray respectively.

analysis results, suggesting that the processing of phrase boundaries selectively activates more cortices, namely the superior temporal cortex, middle temporal cortex, medial orbitofrontal cortex, rostral middle frontal cortex, and rostral anterior cingulate cortex, before the occurrence of the stimuli.

Note-onset-related ERPs versus phrase-boundary-related ERPs. Given that both note onsets and phrase boundaries elicited robust evoked responses, our next goal was to determine whether the brain processes these two stimuli differently by computing the ERP waveform (Fig. 8) and analyzing it with the cluster-based permutation test. The p -value statistics of significant temporal clusters are provided in Table 3.

Figure 9 shows that the two ERPs were significantly different at around -150 , 200 to 450 ms relative to the note onset and phrase markers in both frontal and temporal brain regions, with at least eleven sessions showing significant differences. Figure 10 further illustrates that the differences were mainly localized to the superior temporal cortex followed by the medial orbitofrontal cortex, rostral middle frontal cortex, and rostral anterior cingulate cortex. We also observed significant differences at -150 , 100 to 200 and 400 to 500 ms relative to the stimulus markers in at least six sessions in the caudal middle frontal cortex, insular cortex, and superior frontal cortex (Fig. 11).

Discussion

By integrating the results of within-session analysis, we examined whether note onsets and phrase boundaries elicited different neural responses across subjects. Several temporal clusters of significant difference were identified in the permutation test, demonstrating the difference between the neural responses to note onsets and phrase boundaries in terms of peak lag and amplitude. In addition to the auditory cortex, we were motivated to examine neural responses in frontal brain regions linked to grammatical structure building in studies of speech perception. Besides, the contrast between neural response to note-onsets and phrase boundaries in frontal brain regions may also reflect a process of building up syntactic structures with increasing hierarchy in music, similar to the computation merge in linguistics demonstrated by Zaccarella et al.²⁰

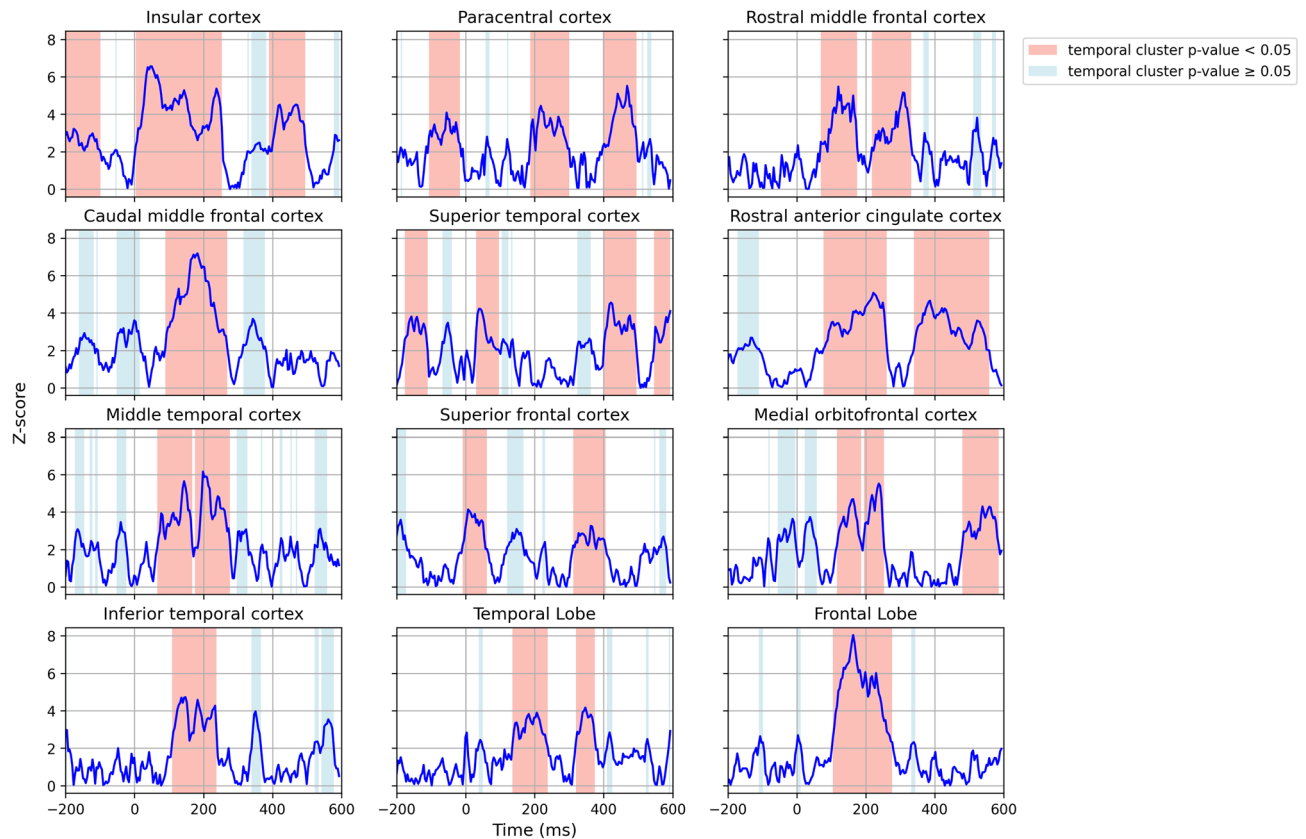


Figure 3. A cluster-based permutation test for each sub-region within a single session (Subject 3, Session 2). The blue line represents the z-score at each time point. The consecutive temporal clusters with z-score > 1.96 are highlighted. The temporal clusters showing significant ($p < 0.05$) difference between the iEEG windows sampled around note onsets and the reference windows are highlighted in red; the non-significant clusters are highlighted in blue.

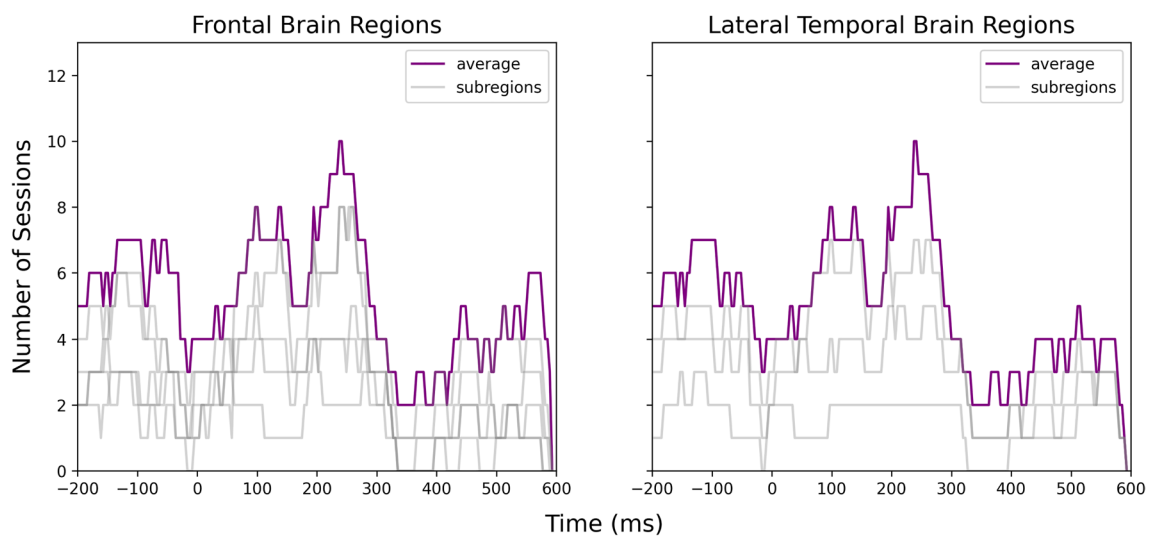


Figure 4. Number of sessions showing significant differences in the iEEG windows sampled around note onsets and the reference windows at each time point. The purple lines were obtained by averaging all channels within the frontal and lateral temporal brain regions. The light gray lines were obtained by averaging the channels within sub-regions that contributed to the frontal and lateral temporal regions.

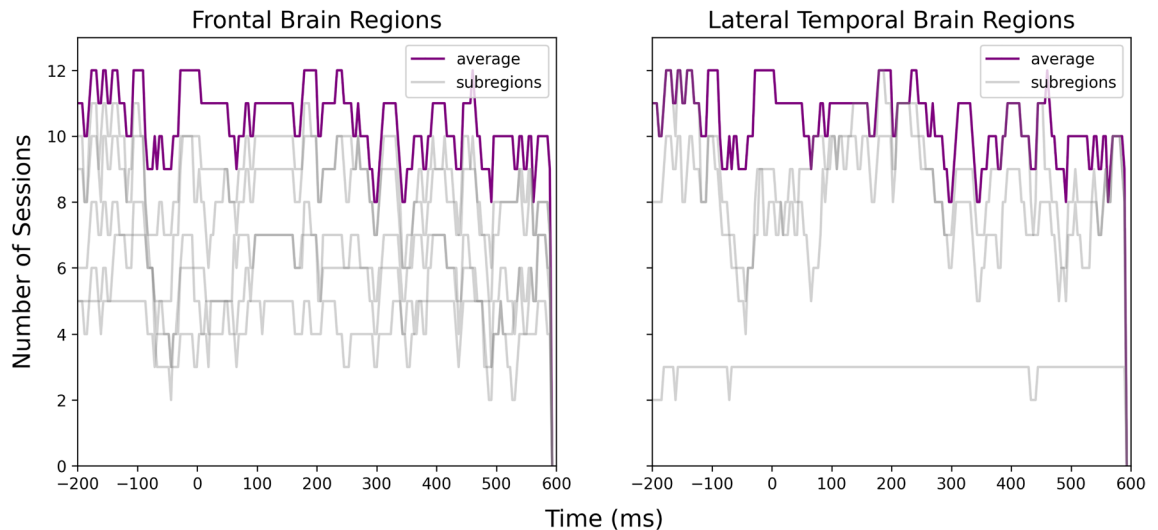


Figure 5. Number of sessions showing significant differences in the iEEG windows sampled around phrase boundaries and the reference windows at each time point. The purple lines were obtained by averaging all channels within the frontal and lateral temporal brain regions. The light gray lines were obtained by averaging the channels within sub-regions that contributed to the frontal and lateral temporal regions.

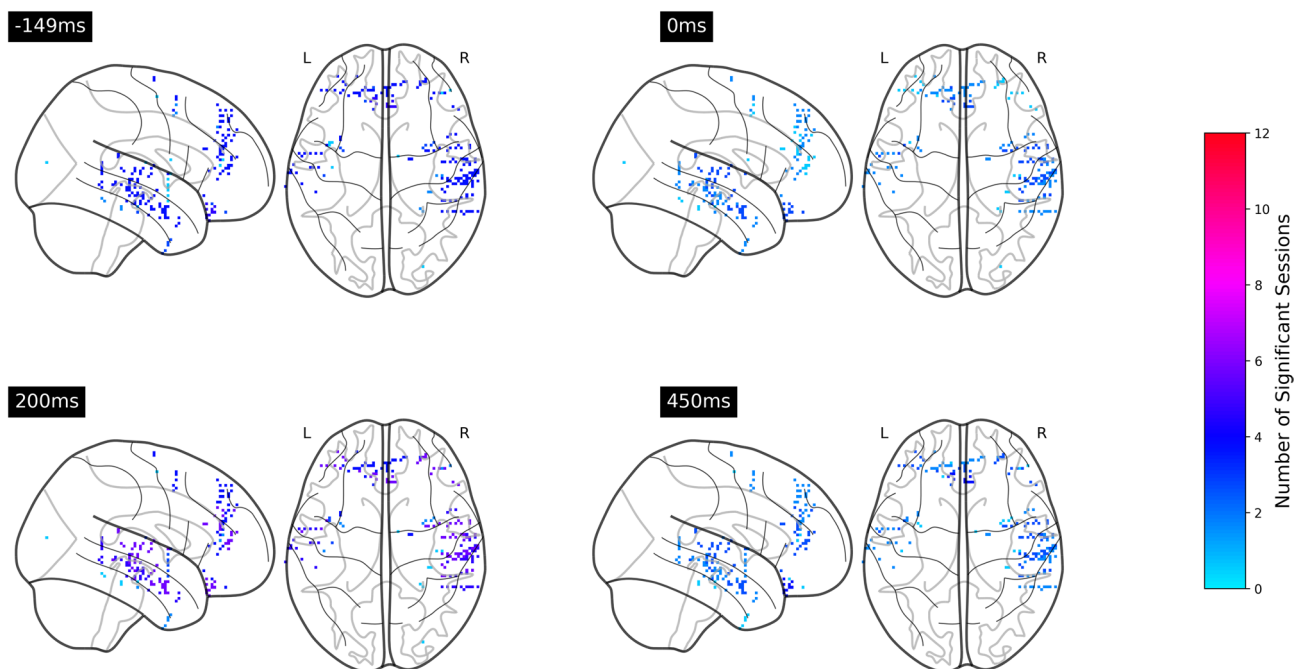


Figure 6. Number of significant sessions within each cortex at -149 , 0 , 200 , 450 ms relative to the note onsets. The electrodes were mapped onto a common MNI space. In the integration analysis, although the location of electrodes differed across sessions, all the electrodes within a cortex have identical color for the purpose of visualization.

We first confirmed that both note onsets and phrase boundaries elicited evoked responses by observing significant statistical differences between the iEEG windows sampled near the stimulus markers and the reference windows sampled between stimulus markers. Although we are the first intracranial EEG study to examine the evoked responses to note onsets and phrase boundaries using the cluster-level statistical analysis, our findings paralleled those of previous ERP studies using averaging techniques. A component was identified around 100 ms and 200 ms after the stimulus onset in both note-onset-related ERPs and phrase-boundary-related ERPs, which resembles the N1-P2 response in the auditory evoked response in language and music²¹⁻²⁴. The N1-P2 like effect suggests that the processing of local cues takes place very quickly after the onset of the stimulus.

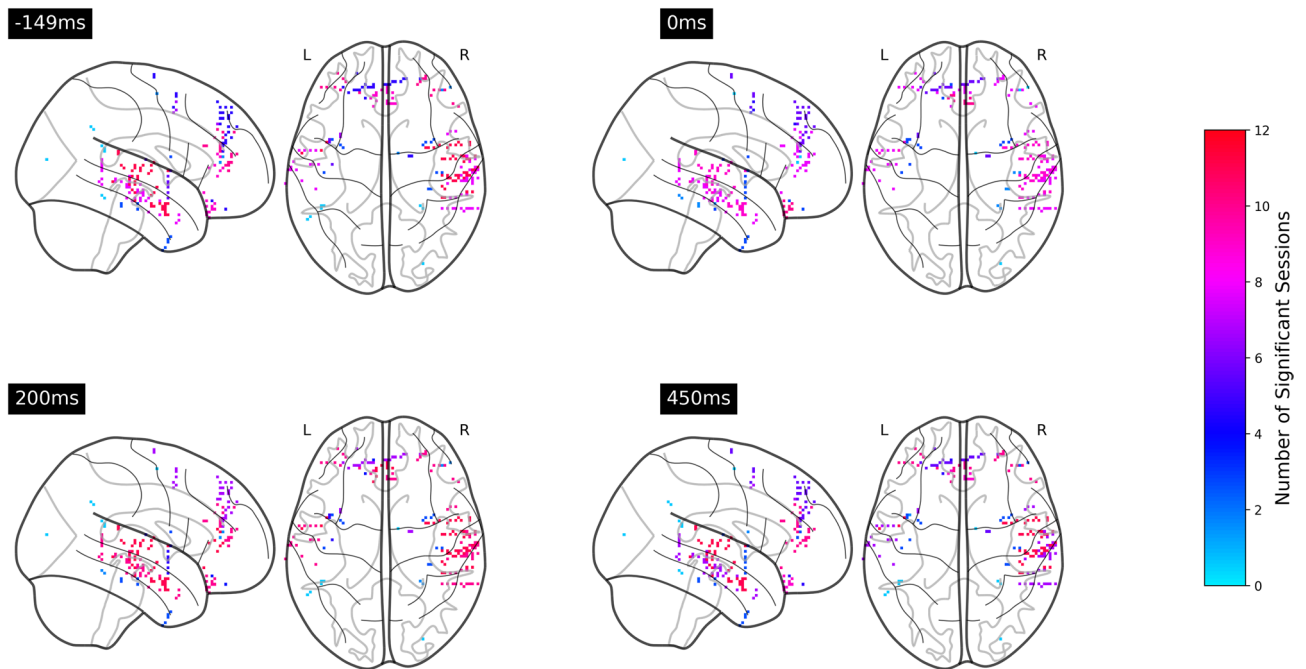


Figure 7. Number of significant sessions within each cortex at -149 , 0 , 200 , 450 ms relative to phrase boundaries.

Subregion	Number of temporal clusters	Mean of p -value	SD of p -value
Middle temporal cortex	53	0.012	0.014
Superior temporal cortex	49	0.010	0.01
Rostral anterior cingulate cortex	58	0.012	0.013
Rostral middle frontal cortex	63	0.011	0.013
Medial orbitofrontal cortex	40	0.012	0.014
Superior frontal cortex	61	0.017	0.014
Insular cortex	53	0.012	0.014
Caudal middle frontal cortex	55	0.009	0.010

Table 3. Statistics of significant temporal clusters in the comparison between iEEG windows sampled around phrase boundaries and iEEG windows sampled around note onsets within each subregion in all twelve sessions.

We then compared the neural responses elicited by note onsets and phrase boundaries and identified three temporal clusters at -150 , 200 , and after 400 ms relative to the stimulus markers with significant differences in at least eleven sessions.

An activation elicited by phrase boundaries at -150 ms was observed in both frontal and lateral temporal brain regions in all twelve sessions. The superior temporal cortex shows structural sensitivity in all twelve sessions, followed by the rostral middle frontal cortex, rostral anterior cingulate cortex, medial orbitofrontal cortex, and middle temporal cortex showing sensitivity in more than ten sessions. Although this prestimulus effect is non-significant in note-onset-related ERPs, we observed an activation of the rostral middle frontal cortex and medial orbitofrontal cortex in at least six sessions, which may reflect the entrainment of cortical rhythm to rhythm of the stimuli²⁵. However, our analysis shows that this component is not consistent across sessions. The -150 ms component unique to the processing of higher-order structures was overlooked in earlier studies of phrase boundaries. The prestimulus activation in the superior temporal cortex could be interpreted by auditory attention, indicating the initiation of a new phrase which does not fit within the expectation of ongoing phrases²⁶. The activation in frontal brain regions suggests a prediction response²⁷, such as a reward effect of positive emotions resulting from anticipatory success. During exposure to music, participants gradually learned the information dynamics and were able to predict forthcoming phrase boundaries, due to changes in note density, melodic themes, key, tempo, and rhythm. This suggests that those neural representations which lead to correct predictions are strengthened and reused. This finding is in line with our previous study on the same dataset²⁸, that an increased frontal theta power was observed during transitions from prolonged musical segments of Mozart's K448 after at least 30-s exposure. The successful prediction of phrase boundaries may preferentially modulate

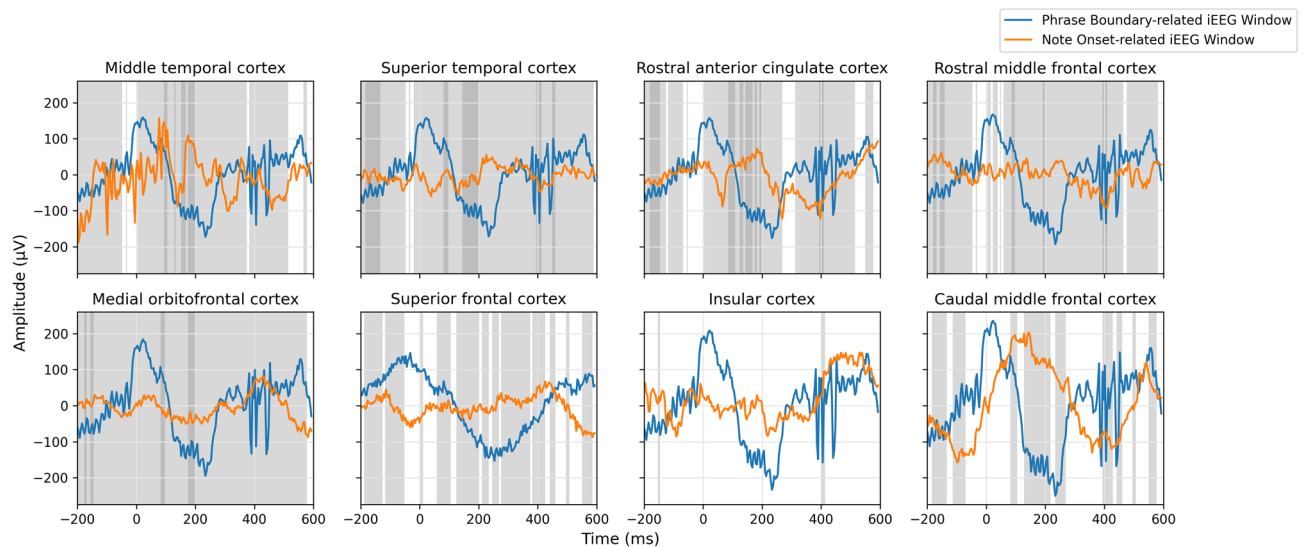


Figure 8. Note-Onset-related and Phrase-Boundary-related ERPs observed at these subregions with significant differences in at least six sessions are shaded in light gray, and ten sessions in dark gray respectively. Despite all note-onset-related iEEG windows being used in the analysis, eight windows were randomly sampled for display due to the large difference between the number of windows sampled around note onset and phrase boundaries and data variance.

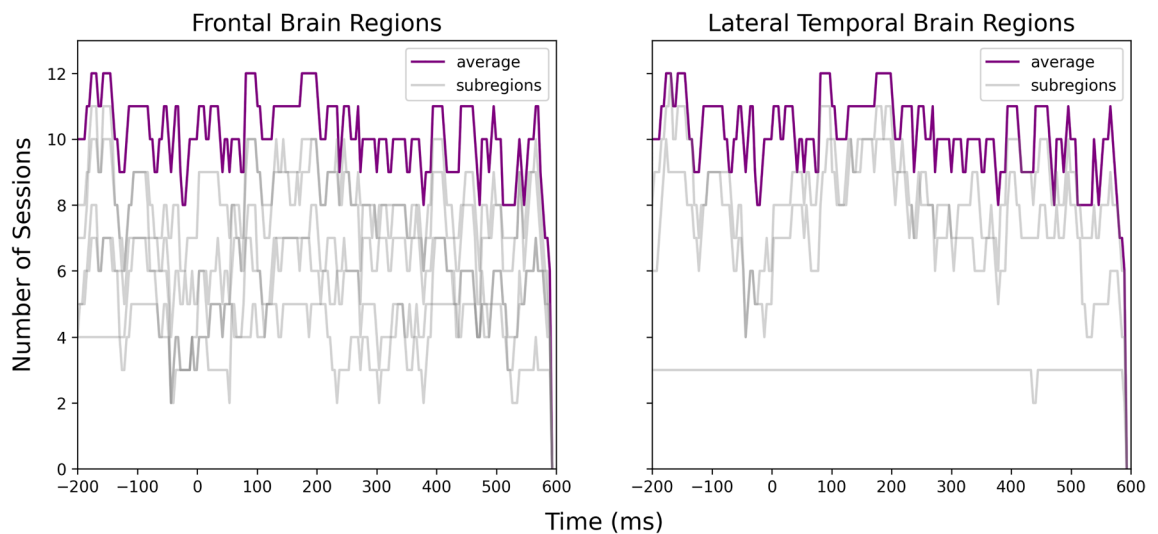


Figure 9. Number of sessions showing significant differences in the iEEG windows sampled around phrase boundaries and note onsets at each time point. The purple lines were obtained by averaging all channels within the frontal and lateral temporal brain regions. The light gray lines were obtained by averaging the channels within sub-regions that contributed to the frontal and lateral temporal regions.

activity in frontal emotional networks, suggesting that the widely observed strong pleasurable responses^{29–31} are linked to the prediction of higher-order musical structures.

Although N1-P2 like components were observed in both note-onset related ERPs and phrase boundaries related ERPs, the significant contrast between the two components, especially in the superior temporal cortex and middle temporal cortex with all twelve sessions showing significant differences, presumably reflects the processing of local cues mediated by more global expectation at phrase boundaries. The timing of the significant difference in the medial orbitofrontal cortex and rostral anterior cingulate cortex is also in line with an early negative component in frontal brain regions which is linked to the building of the grammatical structure in linguistic^{32–34}.

The ERP components at 400 ms and 500 ms post-stimulus onset were only observed around phrase boundaries, potentially indicating higher-order feature extraction for processing the changes in the harmonic and rhythmic structure of the music. The 400 ms post-stimulus component has a broad scalp distribution, maximal in the superior temporal cortex, and is similar to the N400 response in timing thus possibly suggesting the conceptual processing in music³⁵. However, this component is unlikely to be a music N400 response because we did not observe a clear negative-going wave as shown in the prior music N400 work¹². The 500 ms component

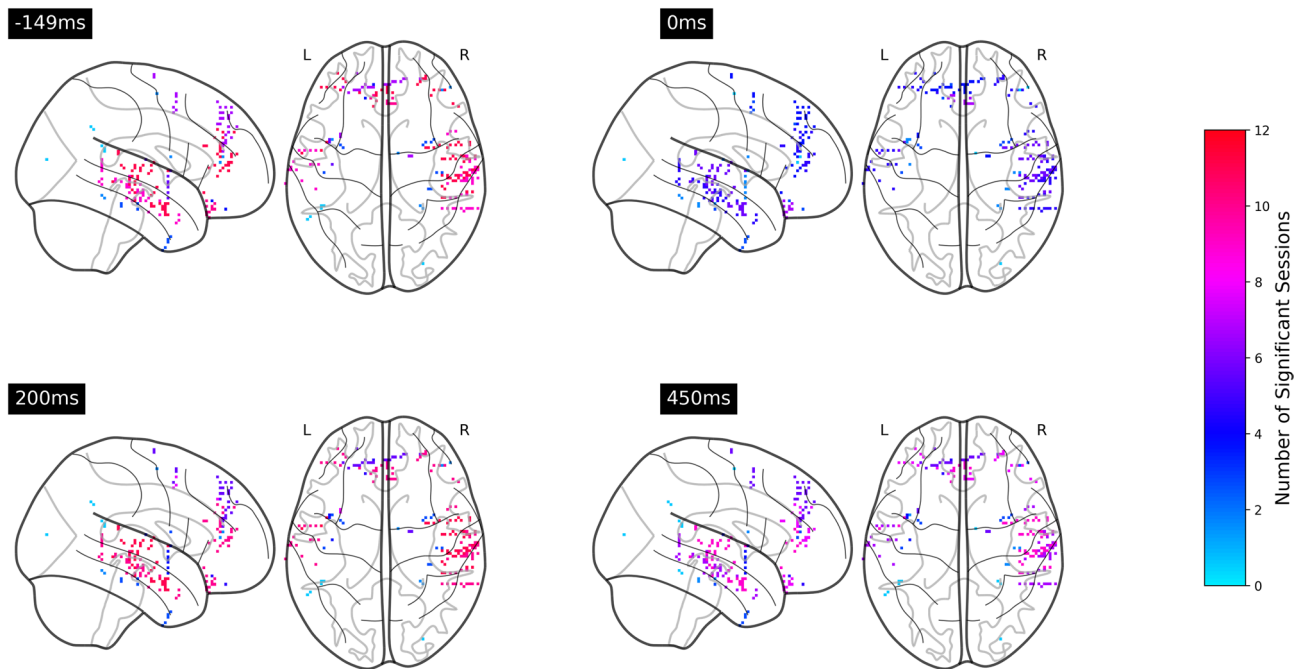


Figure 10. Number of sessions in which neural responses to note onsets and phrase boundaries are significantly different at -149 , 0 , 200 , 450 ms relative to the stimulus markers .

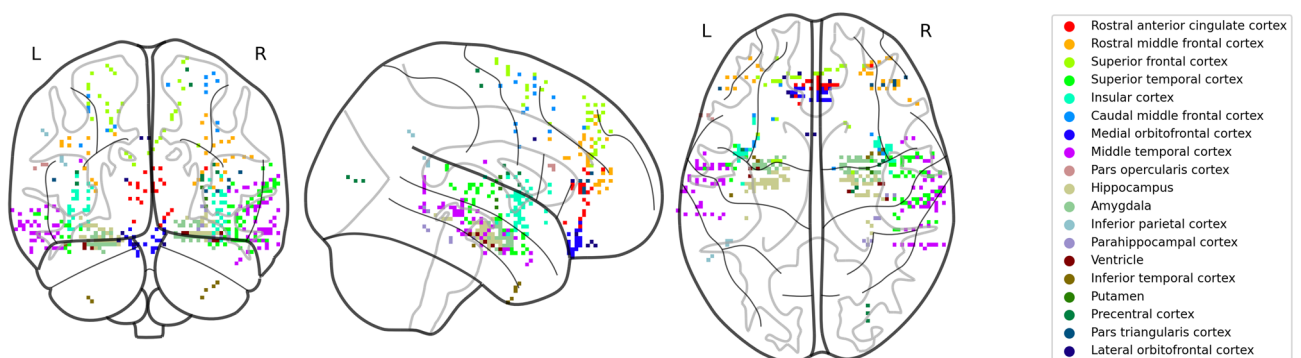


Figure 11. The location of electrodes within each subregion .

resembles CPS discussed in musical phrasing¹⁸. This CPS-like effect was observed in both frontal and temporal brain regions, maximal in the middle temporal cortex and superior temporal cortex. The activation in frontal brain regions suggests that these components may not only reflect the detection of phrase boundaries, but also a violation of melodic expectation in the transition from one phrase to the next. As shown in Fig. 13b, the first 90 s of Mozart's K448 is structurally organized by contrasting melodic themes. The changes at phrase boundaries break the tension built up through harmonic and melodic progression within the previous phrases. Steinbeis et al.³⁶ has reported that a violation of expectation could induce strong emotion. Huron²⁷ further points out that an unexpected but innocuous event may result in anticipatory failure but generate positive emotions, known as the reaction and appraisal responses. Therefore, our findings were in line with the theory of musical expectations and emotion³⁷.

We also analyzed the neural response to note onsets and phrase boundaries in temporal regions as shown in Fig. 12. The posterior temporal regions showed a prestimulus effect on phrase boundaries but not note onsets, which is in line with recent works implicating the sensitivity of these regions in linguistics syntax processing^{38–40}. Besides, the prestimulus effect was observed in posterior temporal regions but not anterior temporal regions, which suggests that this effect is more likely to be induced by music given that the posterior temporal regions are linked to the processing of pitch and temporal variation⁴¹.

The less significant findings of ERPs at note onsets were not unexpected. First, due to the high note density in a naturalistic music excerpt, the iEEG windows sampled around note onsets might cover multiple overlapped ERPs which could not be isolated because the intervals between note onsets were variant. These overlapped ERPs result in the non-significant peaks at -150 and 450 ms. Secondly, the randomly sampled reference windows might also contain ERPs elicited by weak note onsets which were excluded for comparison. To test this

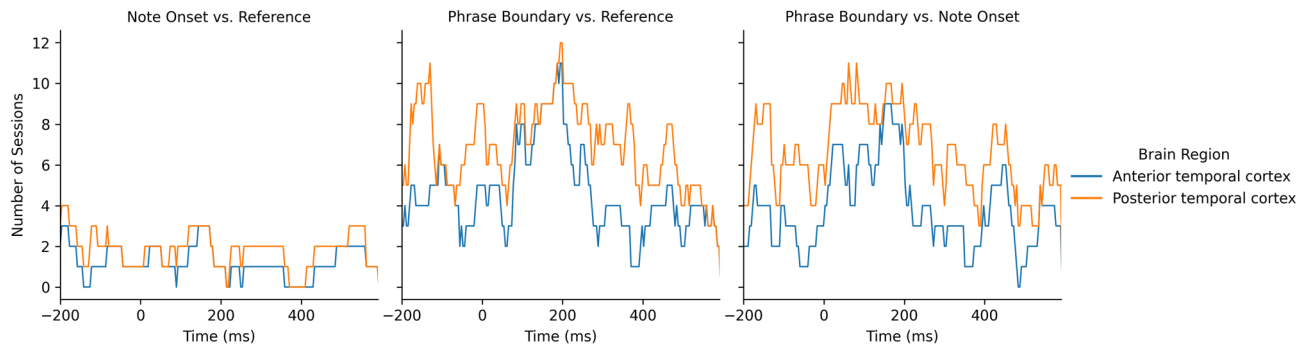


Figure 12. Number of sessions showing significant differences in the iEEG windows sampled in anterior temporal cortex and superior temporal cortex .

hypothesis, we compared the note-onset-related ERPs with iEEG windows randomly sampled during exposure to the silent washed-out period or violet noise. However, the experiment did not yield meaningful results. This might be explained by the yet unknown brain activities that the subjects undergo when not listening to music.

Our analyses extended Quon et al.'s study which shows that the musical structure of K448 may be contributing to its therapeutic effect²⁸ and were performed on the same dataset on which Quon et al. observed a significant interictal epileptiform discharge (IED) reduction in bilateral frontal cortices coupled with increased frontal theta power during transitions from prolonged musical segments after at least 30-s of exposure to K448. It has been reported that listening to specific musical works, such as Mozart's Sonata in D Major for Two Pianos (K448)^{42–44} and the Piano Sonata in C Major (K. 545)⁴⁵, is associated with a reduction in seizure frequency and a reduction in abnormal interictal epileptiform discharges in patients with epilepsy. However, this effect has been demonstrated with only a small number of musical works with similar structures^{46–48}, suggesting that this effect is dependent on musical structures such as a high degree of long-term periodicity^{49,50}. In revealing the potential reward linked to prediction response occurring at phrase boundaries in Mozart's K448, we shed light on the theory that structural organization of Mozart's K448 could explain the mechanism behind music interventions such as the Mozart effect for epilepsy.

The results of our study must be interpreted in light of several limitations. First, we only studied the time-locked evoked and anticipatory responses while music perception also involves oscillatory response which could be estimated by an oscillator model. However, we considered oscillatory response to be trivial in our case because of the interplay between oscillatory and evoked components in auditory processing. Doelling et al.⁵¹ has shown that the evoked response can be reduced by smoothing the attack of note onsets. In contrast, the evoked response is the dominant response to the strong attack of note onsets that we investigated. Another major limitation was the overlapping of multiple note-onset-related ERPs within one window. Most importantly, we would like to acknowledge that the sample size might have limited our ability to generalize our results. The number of subjects was relatively small and 8 phrase boundaries were insufficient compared to 274 note onsets in the same music excerpt. Although previous studies⁵² have shown that 8 trials would be sufficient to detect certain ERP components, the statistical power does not saturate at this number. This could be improved in further studies by introducing more high-order structural changes in longer music excerpts.

In conclusion, our findings demonstrate that musical components at different hierarchical levels in Mozart's K448 evoke consistent differential neural responses. We identify a prestimulus ERP component unique to note onsets occurring at musical phrase boundary, which indicates a predictive response in the frontal brain regions to higher-order structural changes within the music. These findings may guide future investigation of electrophysiological markers for processing hierarchy in music cognition and lead to new insights into potential auditory treatments for neurological disorders such as epilepsy.

Material and methods

Study population. A total of twelve sessions of Intracranial Stereo-EEG data were collected from eight subjects with refractory epilepsy undergoing monitoring for the clinical treatment. The electrodes were implanted based on clinical needs. These subjects had an average normalized baseline IED rate of 1.43 (SD 0.94). Each subject had electrode coverage in both hemispheres with between 34 and 77 artifact-free channels after excluding channels outside of MRI co-registered brain regions and bad channels for which the raw signal was greater than 2.5 standard deviations from the median value across channels. All subjects reported little to no previous musical training and limited exposure to classical music. Other subject demographic and clinical characteristics are provided in Table 4.

All patients provided informed consent to participate in this study, approved by the Committee for the Protection of Human Subjects (CPHS#: 12495) at Dartmouth College. Approval by CPHS was based on the study's appropriate balance of risk and benefit to subjects and a study design in which risks to subjects are minimized. As such, our study followed the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. Specific national laws were also observed, and all details that might disclose the identity of the subjects under study were omitted.

Subject	Gender	Age	Left channels	Right channels	Handedness
1	Female	29	68	0	Right
2	Female	43	0	25	Right
3	Male	30	32	18	Right
4	Male	56	44	45	Right
5	Male	27	51	51	Right
6	Male	27	51	38	Right
7	Male	32	2	38	Right
8	Male	35	60	0	Right

Table 4. Subject demographic. Left channels and right channels denote the number of contacts remaining after the exclusion of bad channels and channels outside of co-registered grey matter regions.

Experiment paradigm. Each session of the experiment lasted approximately 30 minutes, consisting of 9 trials including (1) A baseline period only before the first trial of each session; (2) two minutes of a randomly sampled piece of music. The subject was required to finish the SART attention task, during the last 30 s of the music excerpt to confirm that the subject was attending to the piece of music. The attention task was reported separately; and (3) A washout period of one minute of silence after each music excerpt. Subjects listened to a 90-s violet noise and eight pieces of music including Mozart's Sonata for Two Pianos in D major (K448) during data collection. The trials were repeated in random permutation until each piece of music was presented once. The current data analysis was only performed on sessions in which subjects listened to Mozart's K448. Other auditory stimuli were Frederic Chopin's Bolero in C-Op. 19 for piano, performed by Nikita Magaloff; Franz Liszt's Piano Sonata in B Minor, 1st movement: Lento assai-Allegro energico, performed by Leslie Howard; Wagner's Lohengrin Prelude to Act I; Mozart K448 with boosted 40Hz harmonics; and three songs chosen by each subject from a preferred musical genre (Tumbling Tumble Weeds by Sons of the Pioneers, Barbara Allen by Bradley Kincaid, Jugulator by Judas Priest, Just For by Nickelback, Na Na Hey Hey Kiss Him Goodbye by Steam, Peggy Sue by Buddy Holly). These eight auditory stimuli were excluded due to a lack of ground truth for phrase boundaries.

Stimulus. Figure 13 shows 274 note onsets and 8 phrase boundaries extracted from the music excerpt as low-level and high-level musical components. The note onsets were detected by picking peaks in an onset strength envelope using *librosa*⁵³. To reduce overlapping between iEEG windows sampled around two adjacent note onsets, we excluded 50% of the weak note onsets based on the conclusion of previous studies^{54,55} that increasing stimulus intensity produces an increase in P300 amplitude of the ERP. The phrase boundaries were first annotated by a music expert on the score, and labeled in the audio by aligning the midi generated from the score with the audio using dynamic time warping (DTW). A theoretical evaluation of the first 90 s of Mozart's K448 is performed to analyze the musical structure and annotated on Fig. 13b.

Intracranial stereo-EEG data. iEEG was sampled at 512Hz from either 0.80-mm PMT platinum depth electrodes or 0.86-mm Ad-Tech platinum depth electrodes (Natus Medical Inc.). For all subjects, pre-implant T1-weighted and T2-weighted MRI images were co-registered with postoperative computed tomography (CT) to obtain the position of small-spacing Stereo-EEG depth electrodes. Freesurfer and the Desikan-Killany atlas were used for hippocampal subfield localization and cortical parcellation, and then final electrode positions were manually reviewed by two neuroradiologists⁵⁶⁻⁵⁹. The coordinates of the electrodes were transformed into a common MNI space for display. Figure 11 shows the electrodes placement within each subregion.

Due to the inconsistent electrode coverage across subjects and sessions, all the statistical analysis in this study was performed in a within-session manner.

The data were subsequently notch filtered at 60 Hz and band-pass filtered from 1 Hz to 250 Hz. All data were then re-referenced to an average referential montage, then downsampled to 256 Hz. This study was based on the data collected during exposures to the first 90 s of Mozart's the Sonata for Two Pianos in D major, K448.

Data segmentation. The iEEG data were segmented into windows around the stimuli. Each window started from 200 ms before the stimuli and 600 ms after the stimuli to include all desired ERP components. To generate reference windows for comparison in the analysis of note-onset-related and phrase boundaries-related ERPs, we randomly sampled 800 ms windows between note onsets with as little overlapping as possible. This resulted in 44 reference windows. The iEEG windows were then grouped by cortex and averaged across channels. The number of windows was resampled to 200 for statistical analysis.

IED rejection. We rejected all iEEG windows that contained at least one interictal epileptiform discharge (IED) in at least one channel. The IEDs were detected using a template matching method⁶⁰ which was validated and performed comparably to clinicians and other published detectors⁶¹⁻⁶⁴. Figure 14 shows an example of an IED identified by this detector.

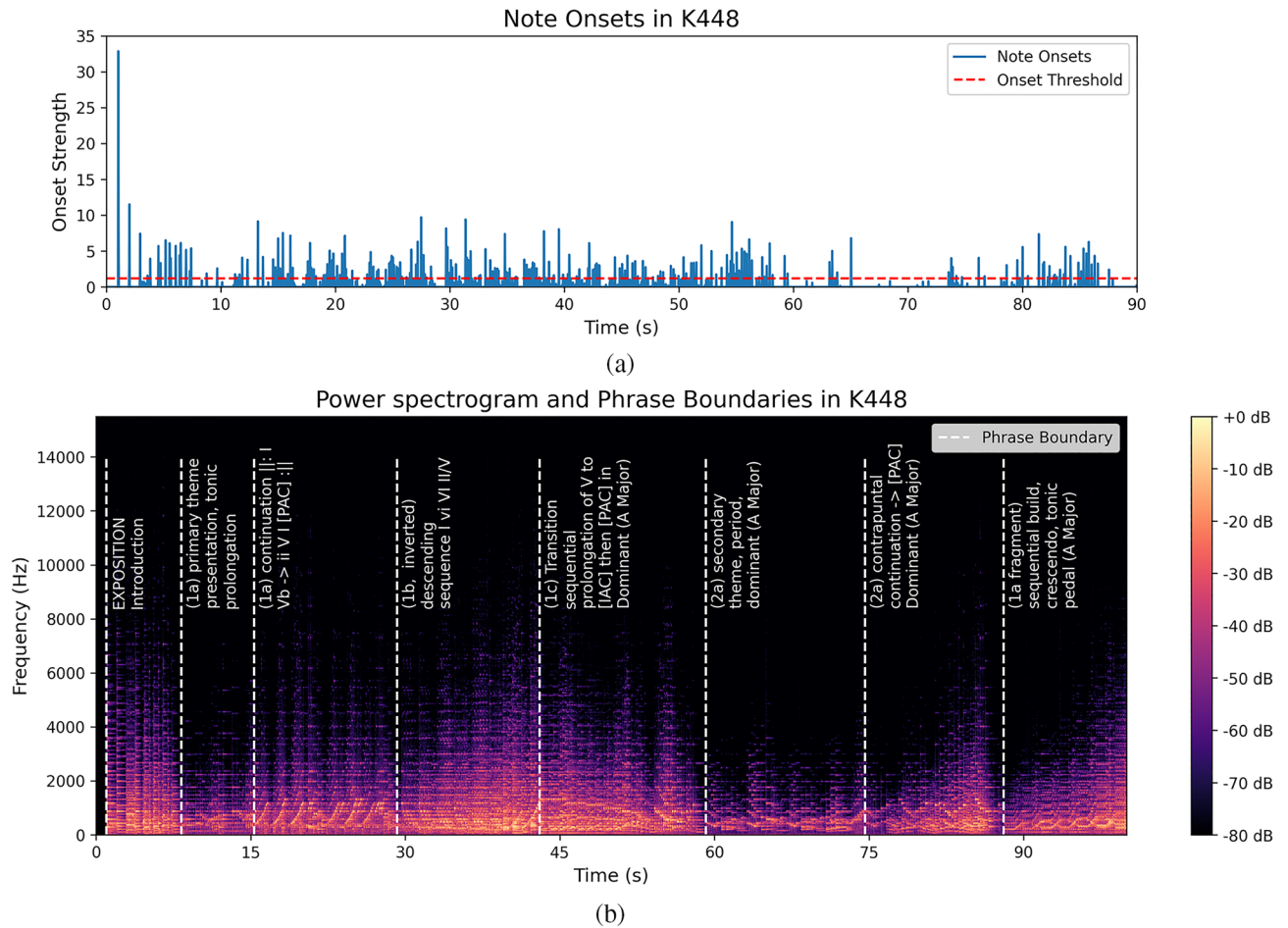


Figure 13. A visualization of the stimuli extracted from the first 90 s of Mozart's K448. In (a), the blue lines mark the normalized note onset strength. 274 note onsets above the threshold marked by the red dotted line are selected. (b) shows the position of 8 phrase boundaries on the spectrogram of first 90 s of K448.

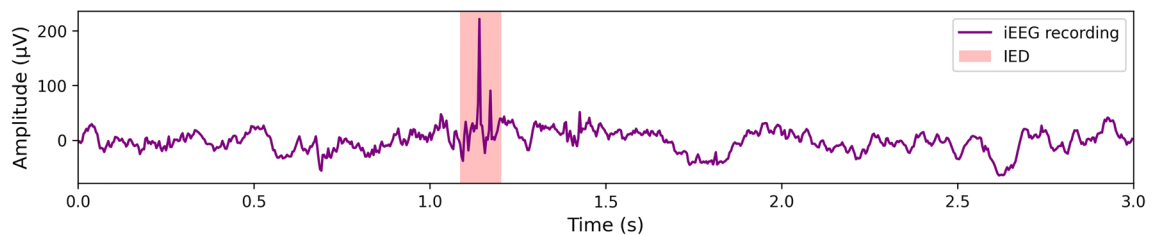


Figure 14. Intracranial recording from an electrode in the rostral anterior cingulate cortex of Subject 1 with an IED highlighted in red.

Statistical analysis. To determine the most appropriate statistical measurement, we first examined the distribution of iEEG signal across channels at each time point within a window by normality test. Different tests were implemented based on the sample size of iEEG windows: D'Agostino and Pearson's test was conducted for iEEG windows sampled around note onsets and reference windows; Shapiro-Wilk test was conducted for windows sampled around phrase boundaries windows. Since the data distribution at each time point was non-Gaussian, a two-sided Mann-Whitney U test was conducted in the statistical analysis.

We utilized a cluster-based permutation test⁶⁵ to identify the consecutive temporal clusters in which neural responses were significantly different in two conditions and thereby verified the existence of note-onset-related and phrase-boundary-related ERPs. The cluster-based permutation test was conducted as follows: (1) A two-sided Mann-Whitney U test was performed at each time point within the window. The U statistics were converted to a time series of Z-scores. (2) The time points with Z-scores larger than the threshold were clustered based on temporal adjacency. The threshold was determined by the Z-score corresponding to the *p*-value of 0.05 in a two-sided test. (3) We repeated steps (1) and (2) on the data permuted for 1000 times if clusters were identified

in step (2). The cluster-level statistics were calculated by taking the maximum of the Z-score within a cluster. The p -value of each cluster was given by the distribution of statistics on the permuted data. (4) We selected the temporal clusters with p -value ≤ 0.05 .

Data availability

Deidentified Stereo-EEG data are available upon reasonable request.

Received: 8 October 2021; Accepted: 25 April 2022

Published online: 10 June 2022

References

- Lerdahl, F. & Jackendoff, R. An overview of hierarchical structure in music. *Music Percept. Interdiscip. J.* **1**, 229–252 (1983).
- Lerdahl, F. & Jackendoff, R. *A Generative Theory of Tonal Music* (MIT Press, 1983).
- Bernstein, L. *The Unanswered Question: Six Talks at Harvard*, vol. Charles Eliot Norton lectures (Harvard University Press, 1976).
- Ding, N., Melloni, L., Zhang, H., Tian, X. & Poeppel, D. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* **19**, 158–164. <https://doi.org/10.1038/nn.4186> (2016).
- Prystauka, Y. & Lewis, A. G. The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Lang Linguist Compass* <https://doi.org/10.1111/lnc3.12347> (2019).
- Bastiaansen, M. & Hagoort, P. Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *J. Cogn. Neurosci.* **27**, 2095–2107. https://doi.org/10.1162/jocn_a_00829 (2015).
- Lewis, A. G., Wang, L. & Bastiaansen, M. Fast oscillatory dynamics during language comprehension: Unification versus maintenance and prediction?. *Brain Lang.* **148**, 51–63. <https://doi.org/10.1016/j.bandl.2015.01.003> (2015).
- Collins, P. & Schmuckler, M. Phrasing influences the recognition of melodies. *Psychon. Bull. Rev.* **4**, 254–9. <https://doi.org/10.3758/BF03209402> (1997).
- Benward, M., Bruce & Saker. *Music: In Theory and Practice*, vol. 1 (McGraw-Hill Higher Education, 2003).
- Chiappe, P. & Schmuckler, M. A. Phrasing influences the recognition of melodies. *Psychon. Bull. Rev.* **4**, 254–259. <https://doi.org/10.3758/BF03209402> (1997).
- Kutas, M. & Hillyard, S. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* **207**, 203–205. <https://doi.org/10.1126/science.7350657> (1980).
- Calma-Rodding, N. & Drury, J. E. Music, language, and the n400: Erp interference patterns across cognitive domains. *Sci. Rep.* **10**, 11222. <https://doi.org/10.1038/s41598-020-66732-0> (2020).
- Miranda, R. A. & Ullman, M. T. Double dissociation between rules and memory in music: An event-related potential study. *Neuroimage* **38**, 331–345. <https://doi.org/10.1016/j.neuroimage.2007.07.034> (2007).
- Osterhout, L. & Holcomb, P. J. Event-related brain potentials elicited by syntactic anomaly. *J. Memory Lang.* **31**, 785–806. [https://doi.org/10.1016/0749-596X\(92\)90039-Z](https://doi.org/10.1016/0749-596X(92)90039-Z) (1992).
- Besson, M. & Faïta, F. An event-related potential (erp) study of musical expectancy: Comparison of musicians with nonmusicians. *J. Exp. Psychol. Hum. Percept. Perf.* **21**, 1278–1296. <https://doi.org/10.1037/0096-1523.21.6.1278> (1995).
- Patel, A. D., Gibson, E., Ratner, J., Besson, M. & Holcomb, P. J. Processing syntactic relations in language and music: An event-related potential study. *J. Cogn. Neurosci.* **10**, 717–733. <https://doi.org/10.1162/089892998563121> (1998).
- Steinhauer, K. & Friederici, A. D. Prosodic boundaries, comma rules, and brain responses: The closure positive shift in erps as a universal marker for prosodic phrasing in listeners and readers. *J. Psycholinguist. Res.* **30**, 267–295. <https://doi.org/10.1023/A:1010443001646> (2001).
- Knösche, T. R. *et al.* Perception of phrase structure in music. *Hum. Brain Mapp.* **24**, 259–273. <https://doi.org/10.1002/hbm.20088> (2005).
- Jongsma, M. L., Desain, P. & Honing, H. Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biol. Psychol.* **66**, 129–152. <https://doi.org/10.1016/j.biopsycho.2003.10.002> (2004).
- Zaccarella, E. & Friederici, A. D. Merge in the human brain: A sub-region based functional investigation in the left pars opercularis. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2015.01818> (2015).
- Sturm, I., Dähne, S., Blankertz, B. & Curio, G. Multi-variate eeg analysis as a novel tool to examine brain responses to naturalistic music stimuli. *PLoS ONE* **10**, 1–30. <https://doi.org/10.1371/journal.pone.0141281> (2015).
- Schaefer, R. S., Desain, P. & Suppes, P. Structural decomposition of eeg signatures of melodic processing. *Biol. Psychol.* **82**, 253–259. <https://doi.org/10.1016/j.biopsycho.2009.08.004> (2009).
- Meyer, M., Baumann, S. & Jancke, L. Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *Neuroimage* **32**, 1510–1523. <https://doi.org/10.1016/j.neuroimage.2006.04.193> (2006).
- Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J. & Ross, B. Modulation of p2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport* **16**, 1781–1785. <https://doi.org/10.1097/01.wnr.0000185017.29316.63> (2005).
- Large, E. W., Herrera, J. A. & Velasco, M. J. Neural networks for beat perception in musical rhythm. *Front. Syst. Neurosci.* **9**, 159. <https://doi.org/10.3389/fnsys.2015.00159> (2015).
- Kaya, E. M. & Elhilali, M. Investigating bottom-up auditory attention. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2014.00327> (2014).
- Huron, D. *Sweet Anticipation: Music and the Psychology of Expectation*, vol. 1. (MIT Press, 2006).
- Quon, R. J. *et al.* Musical components important for the mozart k448 effect in epilepsy. *Sci. Rep.* **11**, 16490. <https://doi.org/10.1038/s41598-021-95922-7> (2021).
- Arjmand, H.-A., Hohagen, J., Paton, B. & Rickard, N. S. Emotional responses to music: Shifts in frontal brain asymmetry mark periods of musical change. *Front. Psychol.* **8**, 2044. <https://doi.org/10.3389/fpsyg.2017.02044> (2017).
- Guhn, M., Hamm, A. & Zentner, M. Physiological and musico-acoustic correlates of the chill response. *Music Percept.* **24**, 473–484. <https://doi.org/10.1525/mp.2007.24.5.473> (2007).
- Grewe, O., Nagel, F., Kopiez, R. & Altenmüller, E. Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Percept.* **24**, 297–314. <https://doi.org/10.1525/mp.2007.24.3.297> (2007).
- Friederici, A. D., Hahne, A. & von Cramon, D. Y. First-pass versus second-pass parsing processes in a wernicke's and a broca's aphasic: Electrophysiological evidence for a double dissociation. *Brain Lang.* **62**, 311–341. <https://doi.org/10.1006/brln.1997.1906> (1998).
- Hagoort, P. & Brown, C. M. Erp effects of listening to speech compared to reading: The p600/sp5 to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia* **38**, 1531–1549. [https://doi.org/10.1016/s0028-3932\(00\)00053-1](https://doi.org/10.1016/s0028-3932(00)00053-1) (2000).
- Hahne, A. & Friederici, A. D. Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *J. Cogn. Neurosci.* **11**, 194–205. <https://doi.org/10.1162/089892999563328> (1999).

35. Daltrozzo, J. & Schön, D. Conceptual processing in music as revealed by N400 effects on words and musical targets. *J. Cogn. Neurosci.* **21**, 1882–1892. <https://doi.org/10.1162/jocn.2009.21113> (2009).
36. Steinbeis, N., Koelsch, S. & Sloboda, J. A. The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *J. Cogn. Neurosci.* **18**, 1380–1393. <https://doi.org/10.1162/jocn.2006.18.8.1380> (2006).
37. Pearce, M. T. & Wiggins, G. A. Auditory expectation: The information dynamics of music perception and cognition. *Top. Cogn. Sci.* **4**, 625–652. <https://doi.org/10.1111/j.1756-8765.2012.01214.x> (2012).
38. Matchin, W., Brodbeck, C., Hammerly, C. & Lau, E. The temporal dynamics of structure and content in sentence comprehension: Evidence from fmri-constrained meg. *Hum. Brain Mapp.* **40**, 663–678. <https://doi.org/10.1002/hbm.24403> (2019).
39. Matchin, W. & Wood, E. Syntax-sensitive regions of the posterior inferior frontal gyrus and the posterior temporal lobe are differentially recruited by production and perception. *Cereb. Cortex Commun.* **1**, tga029. <https://doi.org/10.1093/texcom/tga029> (2020).
40. Matar, S., Dirani, J., Marantz, A. & Pyllkänen, L. Left posterior temporal cortex is sensitive to syntax within conceptually matched arabic expressions. *Sci. Rep.* **11**, 7181. <https://doi.org/10.1038/s41598-021-86474-x> (2021).
41. Liégeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V. & Chauvel, P. Contribution of different cortical areas in the temporal lobes to music processing. *Brain* **121**, 1853–1867. <https://doi.org/10.1093/brain/121.10.1853> (1998).
42. Hughes, J. R., Daaboul, Y., Fino, J. J. & Shaw, G. L. The “mozart effect” on epileptiform activity. *Clin. Electroencephalogr.* **29**, 109–119. <https://doi.org/10.1177/155005949802900301> (1998).
43. Lin, L.-C., Lee, M.-W., Wei, R.-C., Mok, H.-K. & Yang, R.-C. Mozart k.448 listening decreased seizure recurrence and epileptiform discharges in children with first unprovoked seizures: A randomized controlled study. *BMC Compl. Altern. Med.* **14**, 17. <https://doi.org/10.1186/1472-6882-14-17> (2014).
44. Sesso, G. & Sicca, F. Safe and sound: Meta-analyzing the mozart effect on epilepsy. *Clin. Neurophysiol.* **131**, 1610–1620. <https://doi.org/10.1016/j.clinph.2020.03.039> (2020).
45. Govindarajan, R. *et al.* Mozart k.545 mimics mozart k.448 in reducing epileptiform discharges in epileptic children. *Evid.-Based Compl. Altern. Med.* <https://doi.org/10.1155/2012/607517> (2012).
46. Grylls, E., Kinsky, M., Baggott, A., Wabnitz, C. & McLellan, A. Study of the mozart effect in children with epileptic electroencephalograms. *Seizure* **59**, 77–81. <https://doi.org/10.1016/j.seizure.2018.05.006> (2018).
47. Coppola, G. *et al.* Mozart’s music in children with drug-refractory epileptic encephalopathies: Comparison of two protocols. *Epilepsy Behav.* **78**, 100–103. <https://doi.org/10.1016/j.yebeh.2017.09.028> (2018).
48. Hughes, J. R. & Fino, J. J. The mozart effect: Distinctive aspects of the music-a clue to brain coding?. *Clin. Electroencephalogr.* **31**, 94–103. <https://doi.org/10.1177/155005940003100208> (2000).
49. Jenkins, J. S. The mozart effect. *J. R. Soc. Med.* **94**, 170–172. <https://doi.org/10.1177/014107680109400404> (2001).
50. Anderson, W. S., Kudela, P., Weinberg, S., Bergey, G. K. & Franaszczuk, P. J. Phase-dependent stimulation effects on bursting activity in a neural network cortical simulation. *Epilepsy Res.* **84**, 42–55. <https://doi.org/10.1016/j.eplepsyres.2008.12.005> (2009).
51. Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B. & Poeppel, D. An oscillator model better predicts cortical entrainment to music. *Proc. Natl. Acad. Sci. U S A* **116**, 10113–10121. <https://doi.org/10.1073/pnas.1816414116> (2019).
52. Boudewyn, M. A., Luck, S. J., Farrrens, J. L. & Kappenman, E. S. How many trials does it take to get a significant erp effect? It depends. *Psychophysiology* <https://doi.org/10.1111/psyp.13049> (2018).
53. McFee, B. *et al.* librosa: Audio and music signal analysis in python. In *Proceedings of the 14th Python in Science Conference*, vol. 8 (2015).
54. Buchsbaum, M. & Silverman, J. Stimulus intensity control and the cortical evoked response. *Psychosom. Med.* **30**, 12–22. <https://doi.org/10.1097/00006842-196801000-00002> (1968).
55. Sugg, M. J. & Polich, J. P300 from auditory stimuli: Intensity and frequency effects. *Biol. Psychol.* **41**, 255–269. [https://doi.org/10.1016/0301-0511\(95\)05136-8](https://doi.org/10.1016/0301-0511(95)05136-8) (1995).
56. Iglesias, J. E. *et al.* A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution mri: Application to adaptive segmentation of in vivo mri. *Neuroimage* **115**, 117–137. <https://doi.org/10.1016/j.neuroimage.2015.04.042> (2015).
57. Saygin, Z. M. *et al.* High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: Manual segmentation to automatic atlas. *Neuroimage* **155**, 370–382. <https://doi.org/10.1016/j.neuroimage.2017.04.046> (2017).
58. Kwan, P., Schachter, S. C. & Brodie, M. J. Drug-resistant epilepsy. *N. Engl. J. Med.* **365**, 919–926. <https://doi.org/10.1056/NEJMr1004418> (2011).
59. Fischl, B. *et al.* Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355. [https://doi.org/10.1016/s0896-6273\(02\)00569-x](https://doi.org/10.1016/s0896-6273(02)00569-x) (2002).
60. Horak, P. C. *et al.* (2015) Implementation and evaluation of an interictal spike detector. In *Image Reconstruction from Incomplete Data VIII* Vol. 9600 (eds Bones, P. J. *et al.*) 132–142 (International Society for Optics and Photonics, SPIE, 2015). <https://doi.org/10.1117/12.2189248>.
61. Horak, P. C. *et al.* Interictal epileptiform discharges impair word recall in multiple brain areas. *Epilepsia* **58**, 373–380. <https://doi.org/10.1111/epi.13633> (2017).
62. Horak, P. C. *et al.* Implementation and evaluation of an interictal spike detector. In *Image Reconstruction from Incomplete Data VIII* Vol. 9600 (eds Bones, P. J. *et al.*) 132–142 (International Society for Optics and Photonics SPIE, 2015).
63. Janca, R. *et al.* Detection of interictal epileptiform discharges using signal envelope distribution modelling: Application to epileptic and non-epileptic intracranial recordings. *Brain Topogr.* **28**, 172–183. <https://doi.org/10.1007/s10548-014-0379-1> (2015).
64. Quon, R. J. *et al.* Factors correlated with intracranial interictal epileptiform discharges in refractory epilepsy. *Epilepsia* **62**, 481–491. <https://doi.org/10.1111/epi.16792> (2021).
65. Maris, E. & Oostenveld, R. Nonparametric statistical testing of eeg- and meg-data. *J. Neurosci. Methods* **164**, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024> (2007).

Author contributions

R.J.Q., M.A.C., B.C.J. were involved in the conception and design of the study. R.J.Q. collected the data. Y.F. performed the computational analyses and drafted the manuscript. Y.F., R.J.Q., M.A.C., B.C.J. revised and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.A.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022