OXFORD

Cerebral Cortex, April 2020;30: 2600–2614

doi: 10.1093/cercor/bhz263
Advance Access Publication Date: 25 November 2019
Original Article

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

ORIGINAL ARTICLE

Theta and Gamma Bands Encode Acoustic Dynamics over Wide-Ranging Timescales

Xiangbin Teng[1] and David Poeppel[1,2]

[1]Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, 60322 Frankfurt, Germany and [2]Department of Psychology, New York University, New York, NY 10003, USA

Address correspondence to Xiangbin Teng, Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt, Germany. Email: xiangbin.teng@gmail.com

Abstract

Natural sounds contain acoustic dynamics ranging from tens to hundreds of milliseconds. How does the human auditory system encode acoustic information over wide-ranging timescales to achieve sound recognition? Previous work (Teng et al. 2017) demonstrated a temporal coding preference for the theta and gamma ranges, but it remains unclear how acoustic dynamics between these two ranges are coded. Here, we generated artificial sounds with temporal structures over timescales from ~200 to ~30 ms and investigated temporal coding on different timescales. Participants discriminated sounds with temporal structures at different timescales while undergoing magnetoencephalography recording. Although considerable intertrial phase coherence can be induced by acoustic dynamics of all the timescales, classification analyses reveal that the acoustic information of all timescales is preferentially differentiated through the theta and gamma bands, but not through the alpha and beta bands; stimulus reconstruction shows that the acoustic dynamics in the theta and gamma ranges are preferentially coded. We demonstrate that the theta and gamma bands show the generality of temporal coding with comparable capacity. Our findings provide a novel perspective—acoustic information of all timescales is discretised into two discrete temporal chunks for further perceptual analysis.

Key words: asymmetric sampling, discretization, multiplexing, temporal channel, temporal processing

Introduction

Natural sounds contain rich acoustic dynamics over a wide temporal range reflected in their broadband modulation spectra (Nelken et al. 1999; Lewicki 2002; Singh and Theunissen 2003; Narayan et al. 2006), whereas perceptually relevant information often occupies specific timescales. For example, syllabic information in speech unfolds over ~200-ms temporal windows, while phonemic information is conveyed on a timescale of ~50 ms (Rosen 1992). How, then, does the auditory system efficiently extract relevant information to achieve sound recognition? One strategy would be to process acoustic information at every timescale equally and to derive the appropriate perceptual representations by integrating information across all timescales. A different strategy would be to selectively analyze slowly varying auditory attributes, carried over a longer timescale, to guarantee sufficient information for perceptual analysis, and concurrently to extract fast changing dynamics on a shorter timescale to preserve temporal resolution (Poeppel 2003; Boemio et al. 2005; Giraud and Poeppel 2012; Teng et al. 2017). To adjudicate between these alternatives, we test how the auditory system is entrained by acoustic dynamics over wide-ranging timescales.

Phase-locked neural response to sensory stimuli reflects the dynamics of neural populations at different timescales evoked by sensory stimuli and is thought to reveal the corresponding temporal characteristics of sensory processing (VanRullen 2006; Henry and Obleser 2012; Henry et al. 2014; VanRullen et al. 2014). Previous studies used various stimuli, such as speech, music, and amplitude- or frequency-modulated sounds, and found robust phase-locked neural responses in auditory cortical areas below 10 Hz, suggesting a high temporal coding capacity

© The Author(s) 2019. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permission@oup.com.

in the low-frequency range (Luo and Poeppel 2007; Lakatos et al. 2008; Kerlin et al. 2010; Besle et al. 2011; Cogan and Poeppel 2011; Ding and Simon 2012; Henry and Obleser 2012; Kayser et al. 2012; Ng et al. 2012; Wang et al. 2012; Ding and Simon 2013; Herrmann et al. 2013; Lakatos et al. 2013; Peelle et al. 2013; Doelling et al. 2014; Henry et al. 2014; Kayser et al. 2015; Riecke et al. 2015; Zoefel and VanRullen 2015). On the other hand, there is evidence suggesting that the low gamma band plays an important role in syllable processing and comprehension of speech (Palva et al. 2002; Shahin et al. 2009; Kerlin et al. 2010; Peña and Melloni 2011; Morillon et al. 2012; Gross et al. 2013). Previous experiments examined neural oscillatory responses at the low- and high-frequency ranges and demonstrated that both the neural theta and gamma bands, but not the alpha band, robustly track acoustic information (Luo and Poeppel 2012; Teng et al. 2017). However, two key mechanistic questions remain unresolved: how are acoustic dynamics between the theta and gamma ranges tracked or coded by the auditory system? and Is acoustic information coded preferentially with higher precision in the theta band range, given the higher magnitude of neural responses in the low-frequency range?

Here, we test the temporal coding capacity of the human auditory cortex from 4 to 45 Hz by measuring phase-locked neural response. Building on earlier work (Teng et al. 2017), we generated acoustic stimuli with modulation rates centered at theta (4–7 Hz), alpha (8–12 Hz), beta1 (13–20 Hz), beta2 (21–30 Hz), and low gamma bands (31–45) and conducted a match-to-sample task to evaluate listeners' discriminability of different modulation rates. Classification and decoding analyses on the magnetoencephalography (MEG) data were performed to test which neural frequency bands preserve high temporal coding capacity and faithfully encode acoustic dynamics of different timescales. If acoustic information of different timescales can be extracted from the neural signals of specific frequency bands, such results indicate that the auditory system encodes acoustic dynamics by deploying the specific neural bands. We show that the theta and gamma bands encode acoustic information with comparably high precision, while the alpha and beta bands manifest limited temporal resolution. Next, the stimuli modulated at all the timescales can be classified using the theta and gamma bands. Finally, the source localization results show that the neural activity of both the theta and gamma band originates from similar auditory cortical areas.

Methods

Ethics Statement

The study was approved by the New York University Institutional Review Board (IRB# 10–7277) and conducted in conformity with the 45 Code of Federal Regulations (CFR) part 46 and the principles of the Belmont Report.

Participants

Sixteen right-handed participants (nine females; age range: 23–41) took part in the experiment. Handedness was determined using the Edinburgh Handedness Inventory (Oldfield 1971). All participants had normal hearing and reported no neurological deficits. We excluded the data from one participant because of noise issues during neurophysiological recording; therefore, the analysis included the data from 15 participants (nine females; age range: 23–35).

Stimuli

We generated five types of stimuli building on methods used in previous studies (Boemio et al. 2005; Luo and Poeppel 2012; Teng et al. 2017). Each stimulus was 2-s long and generated by concatenating narrow-band frequency-modulated segments, each of which consisted of 100 sinusoids with randomized amplitude, phase, and frequency. The bandwidth of segments was 100 Hz (within a critical band at the center frequencies used). The to-be-concatenated segments for each stimulus type were drawn from a Gaussian distribution with means of 190, 100, 62, 41, and 27 ms and with standard deviations of 30, 15, 6, 4, and 3 ms, respectively. Hence, the distribution of the segment durations aligned with the range of periods of theta (4–7 Hz), alpha (8–12 Hz), beta1 (13–20 Hz), beta2 (20–30 Hz), and low gamma (30–45 Hz) neural bands (Fig. 1A). The frequency-modulated segments could sweep up from 1000 to 1500 Hz or down from 1500 to 1000 Hz. To be concise, hereafter, we refer to the stimulus type with mean segment duration of 190 ms as a "theta (θ) sound," of 100 ms as an "alpha (α) sound," of 62 ms as a "beta1 ($\beta 1$) sound," of 41 ms as a "beta2 ($\beta 2$) sound," and of 27 ms as a "gamma (γ) sound."

For each stimulus type, we generated three samples with different modulation phases. For example, for the stimulus type with a modulation rate in the theta band, the θ sound, we generated three θ sounds with the same modulation rate but different modulation phases. The cochleogram of one example of the each stimulus type (Ellis 2009) and the corresponding prior distribution of segment duration for each stimulus type are illustrated in Figure 1A. Therefore, we have three sounds for each stimulus type and 15 sounds in total for five stimulus types. The 15 sounds were presented repeatedly in the experiment and named "frozen" sounds. In addition, we generated 40 sounds with distinct modulation phases for each stimulus type. Each of the 40 sounds was presented only once and, therefore, named "distinct" sounds, to indicate that each sound for one stimulus type has different modulation phases from the other sounds. In total, we have 200 distinct sounds for five stimulus types.

By introducing the distinct sounds here, we aimed to prevent the participants from performing the behavioral task by memorizing each stimulus. If only the frozen sounds were used here, which were presented repetitively, the participants could possibly memorize each frozen sound and perform the task by comparing memorized frozen sounds, instead of basing judgements on the modulation rate of each sound. Moreover, as the distinct sounds had different modulation phases, this provided baselines for the following analyses on phase-locked neural responses made it possible to conduct stimulus reconstruction from neural signals.

During MEG recording, participants performed a match-to-sample task to differentiate stimulus types (modulation rates), as illustrated in Figure 1B. On each trial, participants were first required to focus on a white fixation cross in the center of a black screen. Then, the screen showed a word in yellow, "sample," and a sample stimulus was presented simultaneously, which was a distinct sound from one of the five stimulus types. After the sample stimulus was over, the screen showed the word "match" and a pair of "match" stimuli selected from the frozen sounds was presented, one of which matched the modulation rate of the sample stimulus. After the second match stimulus was presented, the participants were required to choose by pressing one of two buttons which interval in the match pair matched the sample stimulus. After the response, the next trial started in 1

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

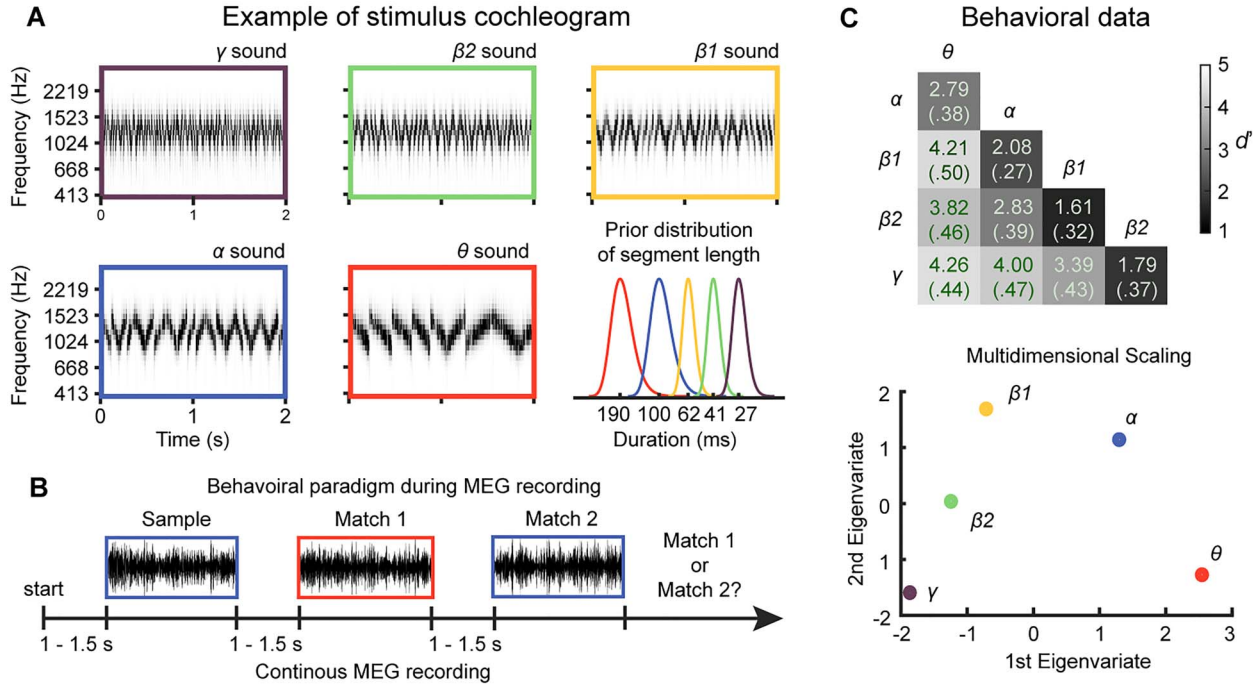


Figure 1. Experimental procedure and behavioral results. (*A*) Cochleograms of five stimulus types. The cochleogram of an example frozen sound from each stimulus type is shown from upper left to bottom middle: θ sound (modulation rate of 4–7 Hz), α sound (8–12 Hz), $\beta 1$ sound (13–20 Hz), $\beta 2$ sound (21–30 Hz), and γ sound (31–45 Hz). Their prior distributions of segment durations are shown in the bottom right panel. The color scheme codes for each stimulus type and is used consistently in the following figures. (*B*) Behavioral paradigm during MEG recording. Participants performed a match-to-sample task to differentiate different stimulus types (modulation rates). The distinct sounds were presented only in the sample interval and the frozen sounds in the match intervals. (*C*) Behavioral results. Upper panel shows group-averaged *d*-prime values in a form of confusion matrix of different pairs of stimulus types. The gray scale codes for *d*-prime values; the number in each cell represents group mean and the standard error of mean across participants (in parentheses). The results of multidimensional scaling on the group-averaged *d*-prime values are plotted in the lower panel and illustrate perceptual distance between different stimulus types. Data points represent each stimulus type.

~ 1.5 s. The intervals between all three stimuli (sample stimulus and two match stimuli) were uniformly distributed between 1 and 1.5 s.

During the match-to-sample task, in the sample intervals, we presented 40 distinct sounds of each stimulus type; in the match intervals, two frozen sounds of each stimulus type were presented 27 times and one frozen sound 26 times. For each pair of stimulus types in comparison, 40 trials were presented to test listeners' discriminability, with 20 trials having one stimulus type of the pair as the sample stimulus and 20 trials having the other stimulus type as the sample stimulus. In total, 200 trials were presented, which contained 200 (5 stimulus types × 40) distinct sounds as the sample stimuli and 400 (5 stimulus types × (27 + 27 + 26)) frozen sounds as the match stimuli.

All the stimuli were normalized to ~65 dB SPL and delivered through plastic air tubes connected to foam ear pieces (E-A-R Tone Gold 3A Insert earphones, Aearo Technologies Auditory Systems).

MEG Recording and Channel Selection

MEG signals were measured with participants in a supine position and in a magnetically shielded room using a 157-channel whole-head axial gradiometer system (Kanazawa Institute of Technology, Japan). A sampling rate of 1000 Hz was used with an online 1–200 Hz analog band-pass filter and a notch filter centered around 60 Hz. After the main experiment, participants were presented with 1-kHz tone beeps of 50-ms duration as a localizer to determine their M100 evoked responses, which is a canonical auditory response (Roberts et al. 2000). Twenty channels with the largest M100 response of both hemispheres (10 channels in each hemisphere) were selected as auditory channels for each participant individually.

Behavioral Data Analysis

Behavioral data analysis was conducted in MATLAB using the Palamedes toolbox 1.5.1 (Prins and Kingdom 2009). For each pair of stimulus types, we averaged correct responses and then converted the percentage correct to d' assuming an independent observer model and an unbiased observer. To avoid infinite *d*-prime values, a half artificial incorrect trial was added in the case that all trials were correct; a half artificial correct trial was added in the case that all trials were incorrect (Macmillan and Creelman 2004).

MEG Data Preprocessing and Analysis

The MEG data analysis was conducted in MATLAB using the Fieldtrip toolbox 20 170 830 (Oostenveld et al. 2011) and wavelet toolbox. Raw MEG data were noise-reduced offline using the time-shifted principle component analysis (de Cheveigné and Simon 2007) and sensor noise suppression (de Cheveigné and Simon 2008). Trials were visually inspected, and those with artifacts such as signal jumps and large fluctuations were discarded. An independent component analysis was used to correct for eye blink-, eye movement-, heartbeat-related and system-related artifacts. Twenty trials were included in the analysis for

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

each frozen sound and 30 trials were included in the analysis for distinct sounds of each stimulus type. Each trial was divided into 4-s epochs, with a 1.5-s prestimulus period, and 2.5-s poststimulus period. Baseline was corrected for each trial by subtracting out the mean of the whole 4-s trial before further analysis.

To extract time–frequency information, single-trial data in each MEG channel were transformed using functions of the Morlet wavelets embedded in the Fieldtrip toolbox, with a frequency range from 1 to 60 Hz in steps of 1 Hz. To balance spectral and temporal resolution of the time–frequency transformation, from 1 to 20 Hz, window length increased linearly from 2 to 10 cycles and was then kept constant at 10 cycles above 20 Hz. Phase and power (squared absolute value) were extracted from the wavelet transform output at each time–frequency point.

The "intertrial phase coherence" (ITPC), a measure of consistency of phase-locked neural activity entrained by stimuli across trials, was calculated on each time–frequency point (details as in Lachaux et al. 1999) for 20 trials of each frozen sound to test the phase-locked neural response on each neural band. The formula of ITPC calculation at a time–frequency point is shown below with "n" indicating the trial number, "f" the frequency point, "t" the time point, and "ϕ" phase at each frequency and time point:

$$\mathrm{ITPC}_{t,f} = \left| \sum_{n=1}^{20} e^{\phi n,t,*i} \right| /20. \quad (1)$$

ITPC was also calculated for the first 20 distinct sounds of each stimulus type to provide a baseline for each neural band. Robust phase-locking on each band can be potentially affected by the power profiles of MEG signals, since high ITPC values in a given neural band may result from its high power magnitude rather than from high phase coherence across trials. If ITPCs of each frozen sound over 20 trials show an effect of phase-locked responses in a neural band whereas ITPCs of the distinct sounds do not in the same neural band, we could conclude that the phase-locked neural responses measured by ITPC are not confounded with the power of this neural band. As ITPC is not normally distributed, the rationalized arcsine transform was applied on ITPC values before further analyses and statistic tests on ITPC (Studebaker 1985).

The evoked power response, which reflects phase-locked neural responses, was computed for each frozen sound by applying the time–frequency transform on an averaged temporal response across 20 trials for each frozen sound. Then, the power values were normalized by dividing the mean power value in the baseline range ($-0.7 \sim -0.3$ s) and taking logarithms with base 10, and then was converted into values with unit of decibel by multiplying by 10. The evoked power was also calculated for 20 distinct sounds of each stimulus type, which was used as baselines to determine significant power responses for the frozen sounds.

The induced power response was calculated for 20 distinct sounds of each stimulus type to match the analysis conducted for the frozen sounds. We first took power of each distinct sound after wavelet transform and then averaged the power over 20 distinct sounds. The baseline correction was the same as used in calculations of the evoked power. We only chose to compute induced power response on distinct sounds because the induced power response calculated from the frozen sounds contains evoked response components and cannot be fully differentiated from the evoked power response. As distinct sounds of each stimulus type have different temporal structures (modulation phases), the evoked component can be, theoretically, averaged out. Furthermore, we calculated induced power responses without baseline correction to show whether the raw power spectra are different for different stimulus types, which may bias ITPC estimations.

The ITPC and power data were averaged from 0.3 to 1.8 s poststimulus onset to minimize the effects of stimulus-evoked onsets and offsets, and within five frequency bands for presenting results of topographies of ITPC: theta (4–7 Hz), alpha (8–12 Hz), beta1 (13–20 Hz), beta2 (21–30 Hz), and gamma1 (31–45 Hz). As the ITPC analysis showed prominent phase-locking effects for $\beta 2$ sounds from 40–52 Hz in our later analyses, we further arbitrarily defined the gamma2 band (40–52 Hz). We referred to frequencies above 30 Hz as the gamma band, which includes the gamma1 and gamma2 bands.

All calculations were first conducted in each MEG channel and then averaged across 20 selected auditory channels. Statistical analyses of ITPC and power were conducted separately for the frozen sounds and the distinct sounds using repeated-measures ANOVA (rmANOVA) across stimulus types on each frequency point. When multiple comparisons were performed, to control false positive rate, adjusted False Discovery Rate (FDR) was used (Benjamini and Hochberg 1995; Yekutieli and Benjamini 1999).

Single-Trial Classification

As the ITPC primarily quantifies circular variance across trials, it does not provide a measure on how the phase patterns of neural signals are unique to each stimulus type. We further conducted single-trial classification analysis of frozen sounds for each stimulus type in each neural band to examine how the auditory system encodes temporal information at different timescales through distinct phase patterns of neural signals. The procedure was described in detail in Ng et al. (2013), and similar methods were also used in Luo and Poeppel (2012), Herrmann et al. (2013), Cogan et al. (2011), and Teng et al. (2017). For three frozen sounds of each stimulus type, one trial was left out for each sound, and then, a template was created by averaging across the remaining 19 trials for this sound (the circular mean is used for phase average). Three templates were created, and the distance between each template and each left-out trial of each frozen sound was computed. The circular distance was applied for phase classification by taking the circular mean over time (0.3–1.8 s) and frequencies within each neural band; the l_2 norm of the linear distance was used for power classification. A left-out trial was given the label of one template if the distance between this trial and the template was the smallest among three templates.

A confusion matrix of classification was constructed by carrying out classification for each trial of each frozen sound of each stimulus type on each auditory channel. Then, the classification performance of each neural band on each stimulus type was measured by converting confusion matrices to d': correctly labeling the target frozen sound was counted as a "hit," while labeling the other two frozen sounds as the target frozen sound was counted as a "false alarm"; d' was calculated based on hit rates and false alarm rates and averaged across all auditory channels. An index of classification efficiency using phase and power response of different frequency band was indicated by the mean of d' over three frozen sounds of each stimulus type, which was compared with the total d' of the identification task (Macmillan and Creelman 2004).

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

Permutation Tests on Uniformity across Stimulus Types for ITPC and Single-Trial Classification

To examine which stimulus type yielded the highest ITPC values or the highest classification performance in a frequency range, we conducted a permutation test, in which we shuffled the labels of the frozen sounds. By doing this, we construct a null hypothesis, which is that each of the five stimulus types does not differ from the other four stimulus types. This procedure directly determined whether a stimulus type showed the highest ITPC values or the highest classification performance compared with the other stimulus types. This procedure avoided conducting paired-*t* tests between each pair of stimulus types, whose null hypothesis is that two stimulus types do not differ from each other.

We first shuffled the labels of the frozen sounds for ITPC values or classification performances of each subject to derive a new value for each stimulus type. The shuffled values were then averaged across 15 participants to derive a group mean. As the new group mean for each stimulus type equally likely contained values from all the five stimulus types, the label of the stimulus type of the shuffled data did not specify real stimulus types any more. In another word, all the stimulus types of the shuffled data could be treated as the same label. Through this shuffling process, the ITPC values or classification performances for frozen sounds are randomly grouped into five labels and the shuffled values can be used as baselines for the true ITPC values of each label. We repeated the shuffling procedure 500 times and derived a one-sided alpha level of 0.01 as a threshold of the group mean of ITPC for each label. We then averaged the thresholds of the five labels as a single threshold for ITPC for all the stimulus types. If the ITPC of one stimulus type is above this threshold, we conclude that the ITPC of this stimulus type is significantly larger than the other stimulus types.

Permutation Test to Determine Baselines for Single-Trial Classification

The classification efficiency in each neural band for each stimulus type could be affected by different baselines between neural bands. For example, in the theta neural band, the classification efficiency is high probably because spontaneous power is low here or because low-frequency ranges (long cycles) have lower variability than, for example, in the alpha band. Hence, the differences of the classification efficiency between neural bands could be determined by different baselines, instead of revealing coding capacity of each neural band. To resolve this issue, we generated a baseline using permutation for the classification efficiency of each stimulus type in each neural band.

We conducted the permutation when each template of each frozen sound of each stimulus type was created (see Single-Trial Classification) for each subject in each neural band. For each stimulus type, there were three frozen sounds. Instead of using the trials for each frozen sound to create the corresponding template, we first shuffled the trials used to create templates (57 in total, 19 for each frozen sound) across the three frozen sounds and disrupted the correspondence between the trials and the frozen sounds. Each template was then created from randomly selected 19 trials. We conducted single-trial classification using these new templates and derived a new group mean for each stimulus type in each neural band. We repeated this shuffling procedure 500 times and derived a one-sided alpha level of 0.01 as a threshold of the group mean, or the baseline, for each stimulus type in each neural band.

MEG Source Reconstruction

As high-resolution structural T_1-weighted MRI scans were only acquired for eight participants, we conducted source reconstruction and localized ITPC and classification efficiency for these participants.

Head shape and head position measurements were taken before the MEG recording session. Both head shape and position were used to coregister individual brain models to each subject's head using uniform scaling, translation, and rotation. The source reconstructions were done by estimating the cortically constrained dynamic statistical parametric mapping of the MEG data. The forward solution (magnetic field estimates at each MEG sensor) was estimated from a source space of 5121 activity points with a boundary-element model method. The inverse solution was calculated from the forward solution. Subsequently, we morphed each individual brain to the FreeSurfer average brain (CorTechs Labs Inc.) and then averaged ITPC and classification efficiency results across eight participants.

We conducted time–frequency analysis to extract phase series and computed ITPC using MNE-Python in source space with parameters comparable to the time–frequency analysis described above on sensor space (Gramfort et al. 2014). We computed ITPC for each frozen sound using 20 trials and then averaged ITPCs from all three frozen sounds for each stimulus type. We selectively presented the source reconstructions of ITPC of each stimulus type in the neural bands that showed significant phase-locking effects in the ITPC analysis on the sensor level (θ sounds in the theta band, α sounds in the alpha band, $\beta 1$ sounds in the beta1 band, $\beta 2$ sounds in the beta2 and gamma2 bands, and γ sounds in the gamma1 band). For single-trial classification, phase series were first exported from Python to MATLAB, and classification efficiency for each stimulus type was calculated using the same procedures as in Single-Trial Classification.

Stimulus Reconstruction

To investigate how acoustic information of different timescales is faithfully encoded by each neural band, we reconstructed cochleograms of different stimulus types from each neural band. The underlying hypothesis is that, if a neural band encodes acoustic dynamics characteristic of a stimulus type, this neural band can be used to reconstruct the cochleograms of this stimulus type with high accuracy, while other neural bands that do not encode acoustic dynamics can only aid in reconstructing the stimulus type to a limited extent.

The method used here is to map between cochleograms of stimuli and the MEG signals. A temporal response function (TRF) was derived from the cochleograms of stimuli ($S(t)$ with subscript c indicating cochlear band number) and their corresponding MEG signals ($R(t)$ with subscript b indicating neural band) through ridge regression with a parameter (lambda) to control for overfitting (superscript t indicating transpose operation):

$$\mathrm{TRF}_{c,b} = \left(R_b^t R_b + \lambda I\right)^{-1} R_b^t S_c. \quad (2)$$

Cochleograms were reconstructed from TRF models as:

$$S_c(t) = \mathrm{TRF}_{c,b} * R(t)_b. \quad (3)$$

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

The reconstruction process included two stages: a training stage and a testing stage (illustrated in Fig. 5A). At the training stage, we used 30 distinct sounds of each stimulus type and their corresponding MEG recordings as a training set to derive TRFs and then used 10 trials from each of three frozen sounds as a validation set to determine the optimal lambda which gave the highest reconstruction performance. At the testing stage, we applied the derived TRFs and lambda values to the remaining 10 trials of each of three frozen sounds and reconstructed the cochleograms of the frozen sounds. Each reconstructed cochleogram was compared with its original cochleogram, and then, model performance was measured by computing Pearson correlation (r) between them. Reconstruction performances from the test set for three frozen sounds of one stimulus type were then averaged. We used the distinct sounds as training samples instead of the frozen sounds because three frozen sounds for each stimulus type represent limited variations of acoustic dynamics, while 30 distinct sounds cover a wide range of variations of dynamics for each stimulus type—each distinct sound is different from the other distinct sounds. Therefore, training using the distinct sounds is not biased towards a specific sample. TRFs were calculated using the Multivariate Temporal Response Function Toolbox (Crosse et al. 2016).

The cochleograms for reconstruction were generated using 8, 16, 32, and 64 bands, separately, ranging from 50 to 4000 Hz (methods described above). This frequency range includes most of the spectral energy of our stimuli centered between 1000 and 1500 Hz. MEG signals were decomposed into theta, alpha, beta1, beta2, gamma1, and gamma2 bands using a two-pass Butterworth filter with an order of two embedded in the Fieldtrip toolbox. Each cochlear band was reconstructed individually from each neural band. The model performance was calculated for each cochlear band and then averaged across all cochlear bands.

One concern with regard to the procedure of stimulus reconstruction here is that the TRF model trained on the basis of a neural band may yield a good stimulus reconstruction merely because the frequency of that neural band overlaps with the modulation rate of the stimuli. Therefore, the stimulus reconstruction performance does not necessarily represent how well each distinct sound is reconstructed. To control for this confound, we conducted a permutation test for each stimulus type in each neural band. All the procedures of stimulus reconstruction remained the same in the permutation test, but the pairings in the training set between the distinct sounds and their corresponding MEG responses were shuffled, yielding a new set in which each distinct sound was paired with an MEG response to a different distinct sound. We conducted this permutation test 500 times and derived a permutation threshold with a one-sided alpha level of 0.01 for each stimulus type and each neural band.

Similar to the permutation tests on uniformity for ITPC and single-trial classification, we derived a threshold for each neural band to determine the stimulus type with the highest reconstruction performance in the neural band. We first shuffled the labels for reconstruction performance of the five stimulus types in each neural band for each subject. We created a new data set and derived a new group mean. We repeated the shuffling procedure 500 times and derived a one-sided alpha level of 0.01 as a threshold of the group mean of reconstruction performance for each label. We then averaged the thresholds of the five labels as a single threshold for all the stimulus types. If the reconstruction performance of one stimulus type is above this threshold in a neural band, we conclude that the reconstruction performance of this stimulus type is significantly larger than the other stimulus types.

Results

Behavioral Performance

The participants' discriminability between different modulation rates (θ, α, $\beta1$, $\beta2$, and γ sounds) increases as the difference between two modulation rates becomes larger (Fig. 1C, upper panel). The discriminability between adjacent modulation rates is best at the low-frequency range but decreases as modulation rates increases (Fig. 1C, lower panel). For example, although the difference of modulation rates between θ and α sounds is smaller than between $\beta2$ and γ sounds, θ and α sounds were better differentiated. We selected d-prime values of four pairs of adjacent modulations (θ vs. α; α vs. $\beta1$; $\beta1$ vs. $\beta2$; $\beta2$ vs. γ) and conducted a one-way rmANOVA with difference of modulation rate between two stimulus types as the main factor. We found a significant main effect ($F(3,42) = 5.11$, $P = 0.040$, $\eta_p^2 = 0.267$).

Phase Coherence and Power Responses for Frozen and Distinct Sounds

The ITPC values for each stimulus type (each modulation rate) of the frozen sounds are plotted from 2 to 60 Hz in Figure 2A. The results of ITPC show robust phase-locked neural responses for all the five stimulus types in their corresponding neural bands, with θ and γ sounds in the corresponding bands showing the prominent phase-locking effects (Fig. 2A). Interestingly, an effect of phase tracking is also observed for $\beta2$ sounds in the gamma2 band (40–52 Hz), which we examined further in subsequent analyses. The phase-locking patterns across temporal regimes are mirrored in topographies of ITPC (Fig. 2B), which manifests clear auditory response patterns for θ sounds in the theta band, γ sounds in the gamma1 band, and $\beta2$ sounds in the gamma2 band. The results of evoked power are consistent with ITPC and show robust power responses across frequency and time for θ and γ sounds in the theta band and gamma band, respectively (Fig. 2D). In contrast, ITPC and power results for distinct sounds do not show any effects in different neural bands, and importantly, the power spectra without baseline correction and the induced power spectra are comparable across all stimulus types (Fig. 2E). The detailed analyses for the above results are shown below.

Phase Coherence

We first averaged ITPC values for each stimulus type across five predefined neural bands—theta (4–7 Hz), alpha (8–12 Hz), beta1 (13–20 Hz), beta2 (21–30 Hz), gamma1 (31–45 Hz), and one neural band defined post hoc—gamma2 (40–52 Hz). We conducted a Stimulus type × Hemisphere × Neural band three-way rmANOVA on ITPC. This revealed the main effects of Stimulus type ($F(4,56) = 9.06$, $P < 0.001$, $\eta_p^2 = 0.392$), Neural band ($F(5,70) = 20.92$, $P < 0.001$, $\eta_p^2 = 0.599$), and Hemisphere ($F(1,14) = 8.34$, $P = 0.012$, $\eta_p^2 = 0.373$). The interaction between Stimulus type and Neural band is significant ($F(20,280) = 20.70$, $P < 0.001$, $\eta_p^2 = 0.597$). Although the main effect of Hemisphere is significant with ITPC values moderately higher in the right hemisphere than the left hemisphere (left: 0.2131 ± 0.0028; right: 0.2166 ± 0.0028), no significant interaction effects were found between Hemisphere and Neural band ($F(5,70) = 1.55$, $P = 0.186$, $\eta_p^2 = 0.100$) and between Hemisphere and Stimulus type ($F(4,56) = 1.31$, $P = 0.279$, $\eta_p^2 = 0.085$). Therefore, in the

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

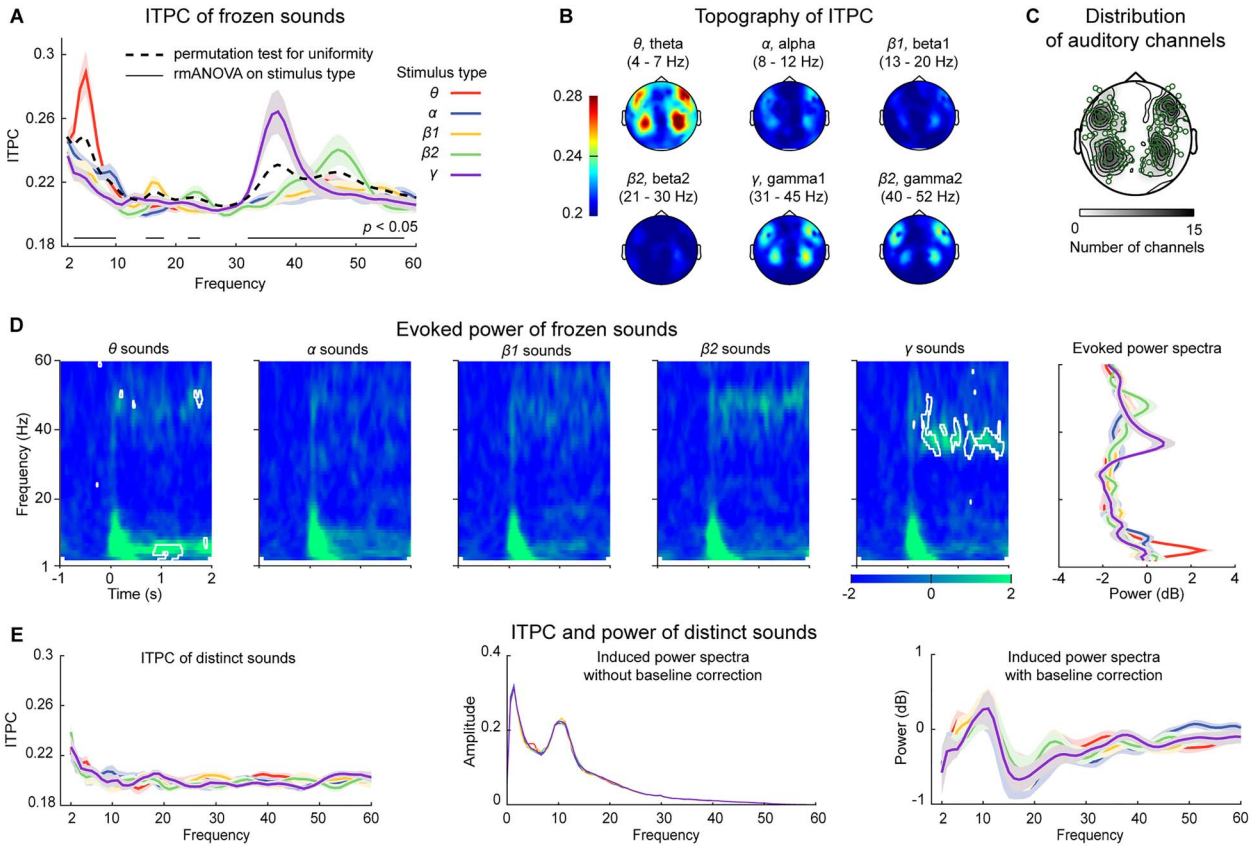


Figure 2. ITPC and power results. (**A**) Spectra of ITPC for the frozen sounds of five stimulus types. The color scheme codes for θ, α, $\beta1$, $\beta2$, and γ sounds, respectively. The shaded areas represent ±1 standard error of the mean across participants. The dashed line represents the permutation threshold for uniformity with a one-sided alpha level of 0.01 (see Methods). ITPC values of the stimulus type above the threshold indicate that the stimulus type evoked significantly higher ITPC values than the other stimulus types. The thin solid line above the x-axis indicates frequencies where the main effect of the stimulus type, tested by rmANOVA, is significant ($P < 0.05$, FDR corrected). The results show significant phase-locked neural responses for all the five stimulus types in their corresponding neural bands, but with the theta band for θ sounds and the gamma band for γ sounds showing the prominent phase-locking effects. (**B**) Topographies of ITPC for each stimulus type. (**C**) A layout of MEG channels that are selected based on the peak of M100 response. Twenty channels are selected for each subject (10 in each hemisphere). The channels selected for analysis are indicated by circles. The contours indicate the extent of overlap of selected auditory channels across participants. (**D**) Evoked power for the frozen sounds. From left to right, each panel shows evoked power responses for θ, α, $\beta1$, $\beta2$, and γ sounds, respectively. The right panel showed the evoked power spectra of the five stimulus types. The white contours indicate spectral-temporal tiles where the evoked power responses of the frozen sounds are significantly larger than the evoked power computed using 20 distinct sounds ($P < 0.05$, FDR corrected). (**E**) ITPC and power results for the distinct sounds. Left panel shows the spectra of ITPC for the distinct sounds for each stimulus type. Middle panel shows induced power spectra without baseline correction. Right panel shows induced power spectra (with baseline correction).

analyses to follow, we combined all the selected auditory channels (Fig. 2C) from both hemispheres.

To investigate the differences of neural oscillatory responses between different stimuli types in different neural bands, we conducted a Stimulus-type one-way rmANOVA on each frequency point of ITPC spectra from 2 to 60 Hz averaged across all auditory channels. This revealed a main effect of Stimulus type ($P < 0.01$) in the frequencies of 2–10, 15–18, 22–24, and 32–58 Hz (Fig. 2A) (FDR correction was applied).

To further examine which stimulus type yielded the highest ITPC values in the significant frequency ranges shown above (2–10, 15–18, 22–24, and 32–58 Hz), we conducted a permutation test and derived a one-sided alpha level of 0.01 as a threshold of the group mean of ITPC for each neural band (see Methods). If the ITPC of one stimulus type is above this threshold, we conclude that the ITPC of this stimulus type is significantly larger than the other stimulus types. We found that, within the frequency ranges that show significant main effects of Stimulus type, ITPC of θ sounds is significantly above the threshold from 3 to 8 Hz, as well as α sounds from 8 to 10 Hz and from 57 to 58 Hz, $\beta1$ sounds from 15 to 18 Hz and from 54 to 56 Hz, $\beta2$ sounds from 22 to 24 Hz and from 43 to 51 Hz, and γ sounds from 32 to 42 Hz. In summary, all stimulus types evoked phase-locked responses in their respective neural bands.

Since high ITPC values can be simply caused by high power in a neural band but not by phase coherence across trials, to provide a baseline of ITPC, we calculated ITPC using the same procedures on 20 distinct sounds for each stimulus type and performed a Stimulus-type one-way rmANOVA on each frequency point of ITPC spectra from 2 to 60 Hz (Fig. 2E, left panel). No significant main effect of Stimulus type was found after FDR correction ($P > 0.05$).

Evoked Power Response

We conducted paired *t*-tests to compare evoked power between the frozen sounds and the distinct sounds. The paired *t*-tests were conducted on each time–frequency point from −0.5 to 1.9 s and from 2 to 60 Hz, and FDR correction was applied. We cut off the power responses from 1.9 to 2 s because at

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

low frequencies (e.g., 2 Hz), large temporal windows used in wavelet analysis did not give a valid estimate of power responses due to the epoch size of each trial. In Figure 2D, we used the white contours to indicate the time–frequency points where significant differences of evoked power between the frozen sounds and the distinct sounds were found. Although salient power responses can be observed in the theta and gamma bands for all stimulus types, significant power responses ($P < 0.05$) entrained by the frozen sounds were mainly found in the theta band for θ sounds and in the gamma1 band for γ sounds.

Induced Power Response for the Distinct Sounds

We calculated induced power with and without baseline correction for the distinct sounds of each stimulus type (Fig. 2E, middle and right panels). The rationale for this analysis was that the components of evoked power could be conceivably averaged out, as the distinct sounds have different modulation phases from each other. Therefore, we could examine, without the influence of time-locked components, how different stimulus types yield different power responses. We performed a Stimulus-type one-way rmANOVA on each frequency point of power spectra and found no significant effects of Stimulus type after FDR correction ($P > 0.05$). This result demonstrates that induced power responses are comparable across the five stimulus types and do not contribute to estimation of phase-locked neural responses.

Classification using Phase and Power for the Frozen Sounds

Classification efficiency of each neural band for each stimulus type was calculated using the MEG signals in both the phase and power domains. If the MEG response from a particular neural band can differentiate three frozen sounds of one stimulus type, this would suggest that this neural band encodes detailed temporal information of this stimulus type (modulation phase of each sound). The results of classification efficiency, surprisingly, demonstrate that the phase information in both the theta and gamma bands reliably differentiated the frozen sounds of all the five stimulus types, while the alpha and beta bands show only moderate classification performance (Fig. 3A). Further analyses show that the frozen sounds of all the stimulus types can be differentiated with comparable accuracy using the phase information of all the frequency bands together (2–60 Hz) (Fig. 3C, left panel) and the classification performance is contributed to primarily by the theta and gamma bands (Fig. 3C, right panel). Together, the results of ITPC, evoked power, and classification show that acoustic dynamics of all timescales used in the present study are encoded mainly by the theta and gamma bands.

We first test whether the phase and power information of each neural band contributes to the differentiation of frozen sounds of each stimulus type by conducting a one-sample *t*-test against zero for classification efficiency of each neural band and each stimulus type. After FDR correction was applied, we found that none of classification efficiencies calculated using power information were significantly above zero for any stimulus types and any neural bands ($P > 0.05$). In contrast, classification efficiencies calculated using phase information were significantly above zero for all the stimulus types and all the neural bands ($P < 0.05$) except for θ sounds in the beta1 band ($t(1,14) = 1.46$, $P = 0.164$), $\beta 1$ sounds in the beta2 band ($t(1,14) = 1.76$, $P = 0.100$), and γ sounds in the alpha band ($t(1,14) = 2.17$, $P = 0.051$). Therefore, in the following analyses, we only investigated classification efficiencies calculated using phase information.

We first conducted a permutation test to determine whether each stimulus type was robustly classified in a neural band (see "Permutation Test to Determine Baselines for Single-Trial Classification" in Methods). The results show that the classification efficiencies for all the stimulus types in all the neural band are above the baselines derived (Fig. 3A). We then conducted a Stimulus type × Neural band two-way rmANOVA. The main effects and the interaction are all significant: Stimulus type ($F(4,56) = 6.00$, $P < 0.001$, $\eta_p^2 = 0.300$); Neural band (Greenhouse-Geisser corrected: $F(5,70) = 12.34$, $P = 0.001$, $\eta_p^2 = 0.468$); Stimulus type × Neural band ($F(20,280) = 9.82$, $P < 0.001$, $\eta_p^2 = 0.412$). To determine whether there are differences of classification performance between neural bands, we did pair *t*-tests to compare different neural bands. We found that the classification efficiencies of the theta band and the gamma band are significantly larger than the alpha, beta1, and beta2 bands but do not differ significantly from each other (Table 1, FDR correction was applied).

Next, to determine for each neural band which stimulus type is preferably encoded and therefore has the highest classification efficiency, we did a permutation test similar to the one that we conducted for testing ITPC values to derive a one-sided alpha level of 0.01 as a threshold (see "Permutation Tests on Uniformity across Stimulus Types for ITPC and Single-Trial Classification" in Methods). If classification efficiency for one stimulus within one neural band is above the derived threshold, we conclude that this stimulus type is significantly better classified than the other stimulus types. We found that θ sounds in the theta band, α sounds in the alpha band, γ sounds in the gamma1 band, and $\beta 2$ sounds in the gamma2 band have classification efficiencies above the permutation thresholds.

One observation from Figure 3A is that the classification efficiencies are much higher for all the stimulus types in the theta and gamma bands compared with the other neural bands. We suspect that the frozen sounds of all the stimulus types are probably classified with comparable efficiencies and are mainly encoded by the theta and gamma bands. To test this, we first computed classification efficiency for each stimulus type using frequencies from 2 to 60 Hz, which included all the neural bands, and conducted paired *t*-tests between each pair of stimulus types. After FDR correction was applied, we found a significant difference of classification efficiency between θ sounds and $\beta 1$ sounds ($t(1,14) = 3.96$, $P = 0.014$), but not between other stimulus types ($P > 0.05$) (Fig. 3C, left panel). This result proves that, although small differences of classification efficiency exist between stimulus types, acoustic dynamics of all the stimulus types are comparably encoded by the neural signals recorded by MEG. This result raises a question: what neural bands are encoding the acoustic dynamics of α, $\beta 1$, and $\beta 2$ sounds if neural oscillatory responses for these three stimulus types are much reduced (Fig. 2A)? We then conducted classification analyses using the phase information of each frequency and show the spectra of classification efficiency (Fig. 3C, right panel), which echoes Figure 3A and shows two elevated regions of classification efficiency within the theta and gamma band ranges.

The classification results demonstrate that acoustic dynamics of all temporal ranges are primarily encoded by the theta and gamma bands, to a comparable degree (Table 1). However, it is worth noticing that the results of ITPC (Fig. 2) did not demonstrate much elevated ITPC values in the theta and gamma bands for all the stimulus types. This could be because ITPC measured circular variance across trials in essence, whereas single-trial classification measured circular mean across trials. We further discussed these results below in Discussion.

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

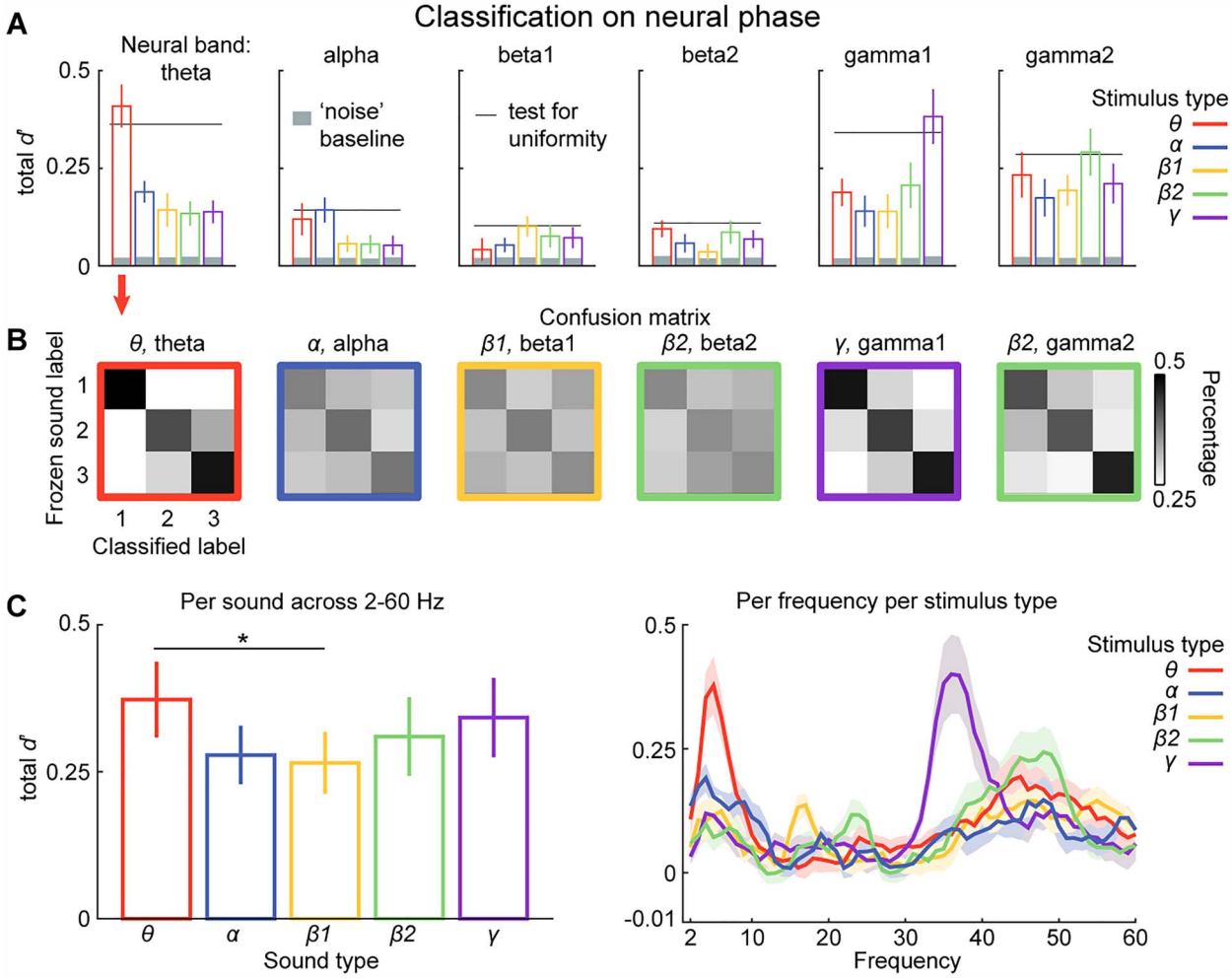


Figure 3. Classification results using the phase information of the MEG response. (**A**) Classification efficiency for each stimulus type within different neural bands. From left to right, each plot shows classification efficiency of each neural band. The dashed line is the permutation threshold for uniformity (see Methods). The shaded areas represent baselines derived from the permutation test (see Methods). Classification performance in the theta, gamma1, and gamma2 bands is significantly higher than in the alpha, beta1, and beta2 bands ($P < 0.05$, see Table 1). (**B**) Group-averaged confusion matrices for each stimulus type in its corresponding bands. The color of the contours of confusion matrices codes for stimulus types. The neural bands represented by each of the confusion matrices are indicated by the arrows and align with the neural bands of **A**. (**C**) Classification efficiencies of full bands and per frequency. Left panel shows classification efficiency for each stimulus type computed using full bands of phase information in the MEG signals and demonstrates that the frozen sounds of all stimulus types are comparably classified. Right panel shows classification efficiency of each stimulus type per frequency. Considerable classification performance can be seen in the theta and gamma ranges for all the stimulus types. This suggests that the major contribution to classification performance of α, $\beta 1$, and $\beta 2$ sounds comes from theta and gamma bands.

Table 1 Paired *t*-test results between classification efficiencies of different neural bands

Neural band	Alpha	Beta1	Beta2	Gamma1	Gamma2
Theta	$t(1,14) = 7.80$, $P < 0.001$	$t(1,14) = 5.36$, $P < 0.001$	$t(1,14) = 5.82$, $P < 0.001$	$t(1,14) = -0.23$, $P = 0.891$	$t(1,14) = -0.41$, $P = 0.792$
Alpha		$t(1,14) = -1.31$, $P = 0.321$	$t(1,14) = -1.55$, $P = 0.242$	$t(1,14) = -3.50$, $P = 0.006$	$t(1,14) = -3.46$, $P = 0.006$
Beta1			$t(1,14) = -0.00$, $P = 0.985$	$t(1,14) = -3.74$, $P = 0.004$	$t(1,14) = -3.80$, $P = 0.004$
Beta2				$t(1,14) = -4.09$, $P = 0.004$	$t(1,14) = -3.90$, $P = 0.004$
Gamma1					$t(1,14) = -1.00$, $P = 0.396$

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

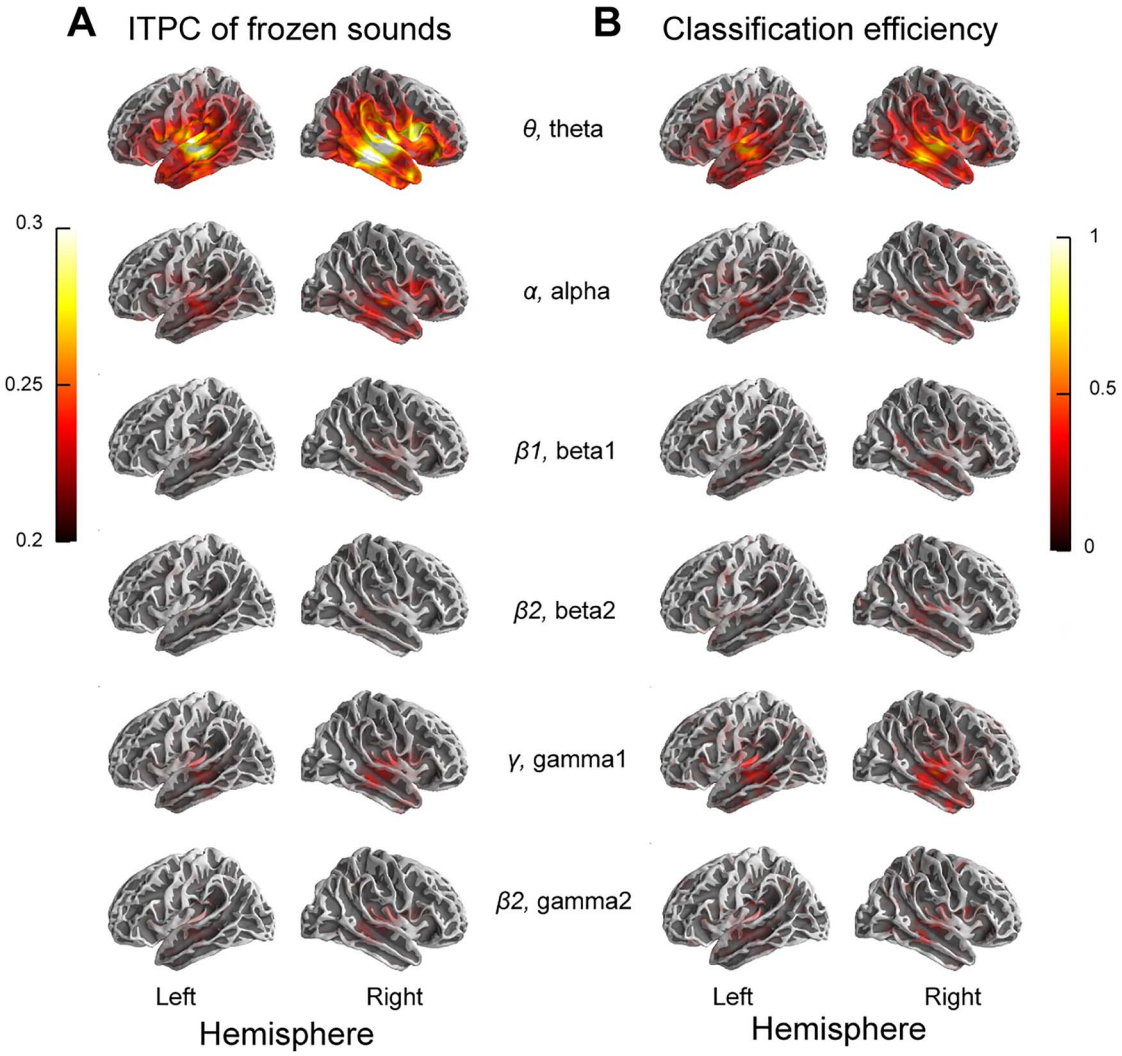


Figure 4. Source localization of ITPC and classification efficiency from eight participants. (**A**) Source plots of ITPC for the frozen sounds. High ITPC values are centered around auditory cortical areas, which demonstrates that robust phase-locked neural responses to different stimulus types originate from similar auditory cortical areas. (**B**) Source plots of classification efficiency for the frozen sounds. Considerable classification performance can be seen for θ sounds in the theta band and γ sounds in the gamma1 band around auditory cortical areas. Moderate classification performance can also be seen for α, $\beta 1$, and $\beta 2$ sounds, but the magnitude is much reduced compared with θ and γ sounds. The legends between **A** and **B** indicate the stimulus type and the neural band. For instance, "θ, theta" indicates that the ITPC values and the classification efficiency on that row represent the results of θ sounds in the theta band.

Source Localization of ITPC and Classification Efficiency

We conducted source reconstruction of MEG signals for eight participants with available MRIs, as high-resolution structural T_1-weighted MRI scans were only acquired for eight participants, and projected the results of ITPC and classification efficiency to source space. Figure 4 shows the source plots of ITPC and classification efficiency averaged across eight participants. High ITPC values for all the stimulus types are centered around auditory cortical areas, and considerable phase-locked neural responses can be observed for all the stimulus types with θ, α, and γ sounds showing higher ITPC values than $\beta 1$ and $\beta 2$ sounds (Fig. 4A). The results of classification efficiency (Fig. 4B), compared with ITPC results, show a different pattern—robust classification performance can only be seen for θ sounds in the theta band and γ sounds in the gamma1 band.

Stimulus Reconstruction

The classification analysis showed that the theta and gamma bands aided in classifying the frozen sounds of different stimulus types (Fig. 3), but this analysis was conducted only on the neural signals, without directly relating the neural signals to the acoustic details of stimuli. Using stimulus reconstruction from the MEG signals, we next examine how neural signals of each band specifically code acoustic information of each stimulus type. The procedures of stimulus reconstruction are shown in Figure 5A. We first reconstructed cochleograms of 16 bands for each stimulus type, to investigate how well each stimulus type can be decoded from different bands of the MEG signals (Fig. 5B). The results show that each stimulus type can be reconstructed sufficiently by its corresponding neural band. We further varied the number of cochlear bands and found that the number of cochlear bands modulates stimulus reconstruction and that θ and γ sounds are reconstructed with comparably high accuracy

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

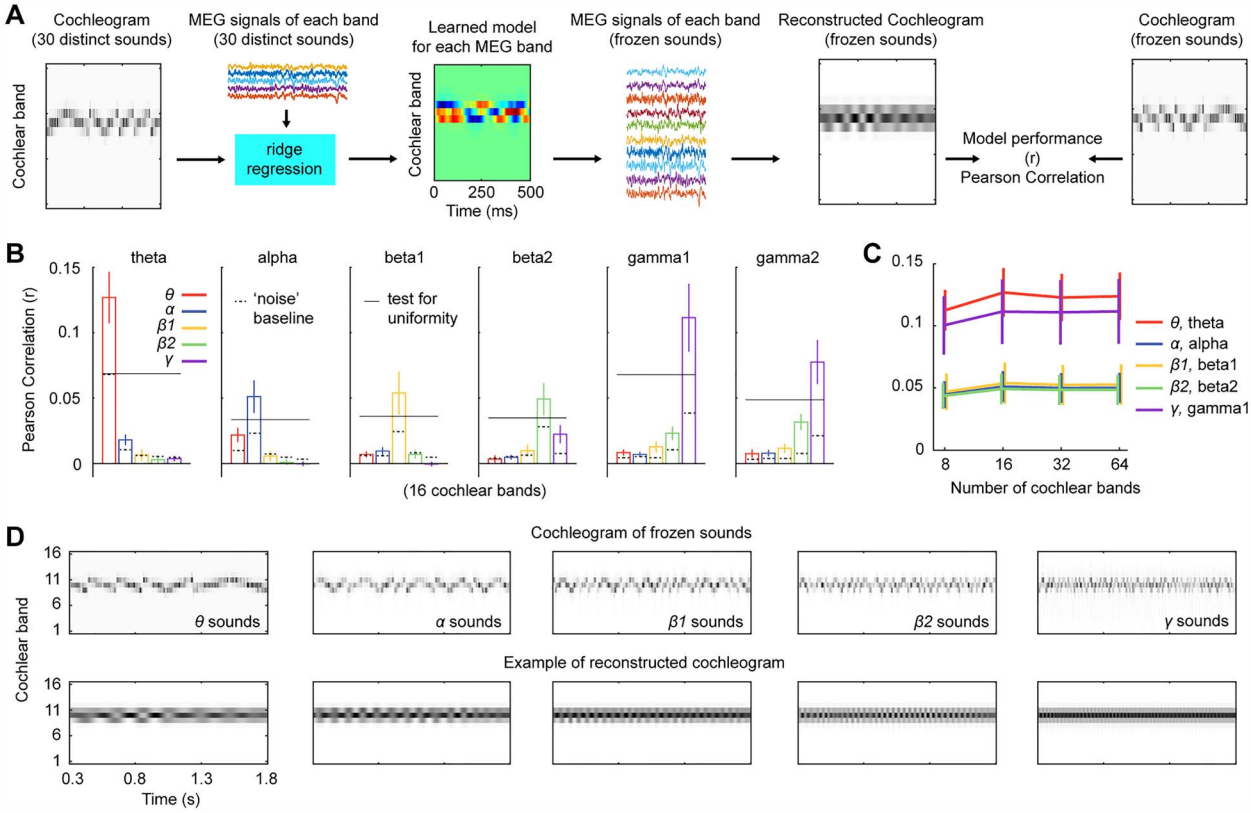


Figure 5. Stimulus reconstruction. (**A**) Illustration of procedures of stimulus reconstruction. Thirty distinct sounds of a stimulus type were used to train a TRF model which was then validated using 10 trials of frozen sounds of this stimulus type to estimate an optimal lambda (see Method for details). The TRF model was tested using the remaining 10 trials of frozen sounds and the model performance was quantified as Pearson correlation between the reconstructed cochleograms of frozen sounds and their original cochleograms. (**B**) Stimulus reconstruction using 16 cochlear bands within each neural band for all the stimulus types. The color scheme codes for different stimulus types; the dashed line within each bar represents the permutation threshold of alpha level of 0.01 for baselines; the thin line in each neural band represents the permutation threshold on the uniformity of alpha level of 0.01. (**C**) Stimulus reconstruction using different numbers of cochlear bands on each stimulus type within its corresponding neural band. The color scheme codes for each stimulus type within its corresponding neural band. The θ and γ sounds can be reconstructed from the respective neural bands with comparably high performance compared with the other stimulus types. (**D**) Examples of reconstructed cochleograms from one subject. From the reconstructed cochleograms, it can be seen that the modulation patterns of different stimulus types are preserved.

from their corresponding neural bands (Fig. 5C). Figure 5D shows examples of reconstructed cochleograms for each stimulus type from one subject.

The performance of stimulus reconstruction using 16 cochlear bands was compared using the threshold of alpha level, 0.01, derived from permutation (see Method). The results (Fig. 5B) show significant reconstruction performance for θ, α, and $\beta1$ sounds in the theta band, θ and α sounds in the alpha band, θ, α, and $\beta1$ sounds in the beta1 band, $\beta1$, $\beta2$, and γ sounds in the beta2 band, and all the stimulus types in the gamma band (gamma1 and gamma2). Similar to permutation tests on uniformity for ITPC and single-trial classification, we derived a threshold for each neural band to determine the stimulus type with the highest reconstruction performance in the neural band. The highest reconstruction performance was observed for each stimulus type in its corresponding band.

We then focused on the reconstruction performance of each stimulus type from its corresponding band using different numbers of cochlear bands. We conducted a Stimulus type $\times$ Cochlear band two-way rmANOVA on reconstruction performance. As reconstruction performance was measured by Pearson correlation, a z-transform on the correlation coefficient was applied before the rmANOVA. We found significant main effects of Stimulus type ($F(4,56) = 9.86$, $P < 0.001$, $\eta_p^2 = 0.413$) and Cochlear band ($F(3,42) = 18.23$, $P < 0.001$, $\eta_p^2 = 0.566$), as well as a significant interaction effect ($F(32,168) = 3.37$, $P < 0.001$, $\eta_p^2 = 0.194$). Post hoc comparisons using paired t-tests with FDR correction on the main effect of Stimulus type show that reconstruction performance of θ and γ sounds is significantly larger than other stimulus types ($P < 0.05$) but not different from each other ($t(1,14) = -0.77$, $P = 0.646$). The linear trend of Cochlear band is significant ($F(1,14) = 17.90$, $P < 0.001$, $\eta_p^2 = 0.561$), suggesting that the decoding performance increases with the number of cochlear bands used in the stimulus reconstruction.

The results of stimulus reconstruction demonstrate that the theta and gamma bands specifically encode the acoustic details of θ and γ sounds, respectively. The reconstruction performance for α, $\beta1$, and $\beta2$ sounds is significant but is much lower than for θ and γ sounds (Fig. 5C).

Discussion

We show that acoustic dynamics are reliably tracked in the theta and gamma bands, consistent with earlier findings (Luo and Poeppel 2012; Wang et al. 2012; Teng et al. 2017). Classification analyses showed that the neural theta and gamma bands contribute to the differentiation of sounds for all stimulus

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

types—with comparable temporal coding capacity (Fig. 3). Source localization results of ITPC and classification efficiency revealed a similar origin of both the neural theta and gamma bands around auditory cortical areas (Fig. 4). Stimulus reconstruction further supported that acoustic dynamics are faithfully encoded by the theta and gamma bands with comparable precision, but only modestly by the alpha and beta bands (Fig. 5). Following previous work, the results demonstrate that the theta and gamma bands code temporal information of all timescales in general and especially extract acoustic features on the timescales of ~30 and ~200 ms. This provides convincing evidence for the hypothesis that the human auditory system primarily analyzes information in two distinct temporal regimes that carry perceptually relevant information (Poeppel 2003). The fact that theta and gamma neural bands can code information of different timescales outside of their frequency ranges shows that the auditory system has a "smart" mechanism to code temporal information, other than the one revealed by neural entrainment or neural following responses—a mapping of one frequency (acoustic) to one frequency (neural). This finding urges researchers to develop a better understanding of a sparse and discrete coding scheme of temporal information in the auditory system. How is acoustic information of different temporal scales transformed and represented on the cortical level of the auditory system? We provide a new perspective—temporal information across all timescales is discretised into two temporal capsules/chunks for further perceptual analysis.

Previous studies on auditory temporal processing typically focused on one temporal regime, mostly in the low-frequency range (<10 Hz). Using various acoustic stimuli, a majority of studies on neural oscillatory responses in auditory cortices suggest a high temporal coding precision primarily in the low-frequency range (Luo and Poeppel 2007; Lakatos et al. 2008; Kerlin et al. 2010; Besle et al. 2011; Cogan and Poeppel 2011; Ding and Simon 2012; Henry and Obleser 2012; Kayser et al. 2012; Ng et al. 2012; Wang et al. 2012; Ding and Simon 2013; Herrmann et al. 2013; Lakatos et al. 2013; Peelle et al. 2013; Doelling et al. 2014; Henry et al. 2014; Kayser et al. 2015; Riecke et al. 2015; Zoefel and VanRullen 2015), several of which indicate that the temporal coding precision of the auditory system decreases with increased modulation rates (Kerlin et al. 2010; Lakatos et al. 2013; Kayser et al. 2015). However, several other studies provide an alternative view. A study using amplitude modulation created by binaural beats showed strong phase-locked neural responses in both the theta and gamma bands (Ross et al. 2014). Recordings in the primary auditory cortex of monkeys also show a phase-locked response using amplitude modulation at 30 Hz (Brosch et al. 2002; Johnson et al. 2012). Gamma band oscillatory response is also found to contribute to speech separation in multiple speaker environments (Kerlin et al. 2010). Seen in conjunction with our own previous work (Luo and Poeppel 2012; Teng et al. 2017), the current findings argue for two concurrent temporal channels for auditory processing, with comparable temporal processing capacity. The auditory system employs the theta and gamma bands to tune to acoustic information of wide-ranging timescales with a preference for the acoustic dynamics in the theta and gamma ranges, which leads to a temporal multiplexing of sensory information (Panzeri et al. 2010; Gross et al. 2013). This may facilitate the efficient extraction of perceptual information at different timescales in speech, such as phonemic scale information and syllabic scale information (Rosen 1992).

Although the theta and gamma neural bands show an effect for all the stimulus types in the classification analysis (Fig. 3), such a generality of temporal coding is not observed in the ITPC (Fig. 2)—ITPC values in the theta and gamma bands were not elevated for all the stimulus types. It could be because ITPC and single-trial classification are quantifying two different aspects of the neural phase series. The ITPC mainly quantified the circular variance of phase across trials, whereas the classification analysis depended on the circular mean—each template in single-trial classification was created by averaging phases across trials. Furthermore, the templates of single-trial classification were created across 150 time points and within a certain bandwidth—a high-dimensional space. Therefore, such a high-dimension space would provide more information for each frozen sound than ITPC at a time–frequency point. This demonstrates that high ITPC does not necessarily suggest distinct neural patterns evoked by acoustic dynamics in certain neural bands, such as the alpha band and the beta band.

The fact that the theta band encodes temporal information of all the timescales here is probably a result of a "chunking" or segmentation process—the auditory system actively chunks sounds into segments of around 150–300 ms, roughly a cycle of the theta band, for grouping acoustic information (Ghitza and Greenberg 2009; Ghitza 2012; Teng et al. 2018). Although different stimulus types have acoustic dynamics on different timescales, this chunking process actively groups acoustic information into chunks in the approximate time window of a theta period. The theta band signals reflect this chunking process and, therefore, can be used to classify all stimulus types (Teng et al. 2018). On the other hand, each gamma cycle integrates fine-grained acoustic information on a local scale (e.g., ~30 ms), for example, transient segment onsets in the stimuli. Therefore, the gamma band reflects encoding temporal information of each stimulus type on a local scale, in comparison with the theta range, and can also be used to classify all the stimulus types (Poeppel 2003).

To further examine the argument arising from previous work (Teng et al. 2017), we hypothesize that the marked reduction of temporal coding in the alpha and beta bands suggests that these two bands play a different processing role in the cortical auditory system. It has been well established that, in the auditory system, the computations implied by the neural alpha band may be related to attention, memory load, listening effort, or functional inhibition (Weisz et al. 2011; Obleser and Weisz 2012; Obleser et al. 2012; Strauß et al. 2014; Wöstmann et al. 2015; Wilsch and Obleser 2016). Such observations have also been shown in the visual and somatosensory systems (van Dijk et al. 2008; Haegens et al. 2011). Beta band neural signals are argued to play a role in predictive coding (Arnal and Giraud 2012; Arnal et al. 2015). Therefore, neural coding schemes in the auditory system may be organized based on timescales to optimize sensory input selection (Buzsáki 2004), with the theta and gamma bands primarily responsible for the temporal coding of acoustic information.

The source localizations of ITPC and classification efficiency demonstrate that the neural theta and gamma bands originate from similar auditory cortical areas, which suggests that the two temporal channels coexist in the cortical auditory system. Although we also found activation for alpha and beta bands around similar cortical areas, the temporal coding precision measured by the classification efficiency is sharply reduced in comparison with the theta and gamma bands. Specifically, in the alpha band, the results from eight participants show the moderate magnitude of ITPC but much reduced classification efficiency (Fig. 4). This suggests that the reduced

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

temporal coding capacity in alpha and beta bands revealed by our analyses is not because MEG fails to record neural activity in the alpha and beta bands, but because the preferred temporal coding in audition is confined to the theta and gamma ranges.

Our finding of robust temporal coding within the theta and gamma ranges is consistent with previous behavioral studies and has fundamental implications. Two perceptual time constants are often found in behavioral studies (Green 1985): experiments on temporal integration converge on a time constant of hundreds of milliseconds (Plomp and Bouman 1959; Green 1960; Zwislocki 1960; Jeffress 1964; Green and Swets 1966; Jeffress 1968; Zwislocki 1969), while studies examining the high temporal resolution of the auditory system show a time constant of less than 30 ms (Viemeister 1979; Forrest 1987; Moore 1988). One recent behavioral study also converges with the present neurophysiological results and demonstrates that the auditory system works concurrently on a short timescale (~30 ms) to extract fine-grained acoustic temporal detail and on a longer timescale to process global acoustic patterns (>200 ms) (Teng et al. 2016). However, the behavioral task in this study, designed to reveal listeners' discriminability between different modulation rates, did not yield results in line with our neurophysiological findings. The behavioral results here show a pattern of a low-pass filter shape—listeners' performance is highest in the low-frequency range and decreases with increased modulation rates, which is consistent with the temporal modulation transfer function found in modulation detection paradigms (Dau et al. 1997). The discrepancy between the current behavioral results of modulation discrimination and our neurophysiological results reveals the complicated nature of auditory temporal processing. As the tasks of detection and differentiation of temporal modulations do not require listeners to decipher information embedded in each modulation cycle, the behavioral results of those tasks probably cannot reflect how the auditory system processes acoustic information on each timescale. Our neurophysiological results invite new behavioral paradigms that target auditory processing on timescales between ~30 and ~200 ms.

Natural sounds contain information at multiple scales and, in order to efficiently sample perceptual information, the auditory system chunks continuous sounds using temporal windows of specific sizes instead of processing acoustic information in a continuous manner (Ghitza and Greenberg 2009; Giraud and Poeppel 2012). Selective representation of acoustic information using timescales of ~30 and ~200 ms may align with efficient encoding—the auditory system preferably extracts acoustic features of the timescales essential to natural sounds (Lewicki 2002; Smith and Lewicki 2006). Such a processing scheme is in line with findings in the visual modality (VanRullen and Koch 2003; VanRullen 2006; Blais et al. 2013), for which preferred encoding on ecologically important features is well demonstrated (Olshausen and Field 2004). One model of auditory processing proposes that—although a very high-resolution is represented in subcortical areas—on the cortical level of the auditory system, there are two main temporal windows used for processing perceptually relevant information: one centered around 200 ms and the other around 30 ms (Poeppel 2003; Giraud and Poeppel 2012). Our results on the theta and gamma bands argue for a segregation of function in the auditory system between low and high processing rates—perhaps optimized for sensory sampling—by an intermediate rate perhaps optimized for allocating attentional and memory resources and functionally inhibiting task- or stimulus-irrelevant actions.

Funding

National Institutes of Health (5R01DC005660 to D.P.); Max-Planck-Society.

Notes

We thank Jeff Walker and Jess Rowland for their technical support. We thank Nina Kazanina, Sean Lee, and Ava Kiai for helpful discussions and comments. *Conflicts of Interest*: The authors declare no competing financial interests.

Author Contributions

X.T. and D.P. designed the study. X.T. performed the experiment, analyzed the data, and wrote the first draft of the paper; X.T. and D.P. wrote the paper.

References

Arnal LH, Doelling KB, Poeppel D. 2015. Delta-Beta coupled oscillations underlie temporal prediction accuracy. *Cereb Cortex*. 25:3077–3085.

Arnal LH, Giraud A-L. 2012. Cortical oscillations and sensory predictions. *Trends Cogn Sci*. 16:390–398.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met*. 57:289–300.

Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE. 2011. Tuning of the human neocortex to the temporal dynamics of attended events. *J Neurosci*. 31:3176–3185.

Blais C, Arguin M, Gosselin F. 2013. Human visual processing oscillates: evidence from a classification image technique. *Cognition*. 128:353–362.

Boemio A, Fromm S, Braun A, Poeppel D. 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci*. 8:389–395.

Brosch M, Budinger E, Scheich H. 2002. Stimulus-related gamma oscillations in primate auditory cortex. *J Neurophysiol*. 87:2715–2725.

Buzsáki G. 2004. Neuronal oscillations in cortical networks. *Science*. 304:1926–1929.

Cogan GB, Poeppel D. 2011. A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *J Neurophysiol*. 106:554–563.

Crosse MJ, Di Liberto GM, Bednar A, Lalor EC. 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. *Front Hum Neurosci*. 10:3958–3914.

Dau T, Kollmeier B, Kohlrausch A. 1997. Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *J Acoust Soc Am*. 102:2906.

de Cheveigné A, Simon JZ. 2007. Denoising based on time-shift PCA. *J Neurosci Meth*. 165:297–305.

de Cheveigné A, Simon JZ. 2008. Sensor noise suppression. *J Neurosci Meth*. 168:195–202.

Ding N, Simon JZ. 2012. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc Natl Acad Sci U S A*. 109:11854–11859.

Ding N, Simon JZ. 2013. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *J Neurosci*. 33:5728–5735.

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

Doelling KB, Arnal LH, Ghitza O, Poeppel D. 2014. Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*. 85(Part 2 IS):761–768.

Ellis DPW. 2009. Gammatone-like spectrograms. *web resource*. http://www.ee.columbia.edu/~dpwe/resources/matlab/gammatonegram/.

Forrest TG. 1987. Detection of partially filled gaps in noise and the temporal modulation transfer function. *J Acoust Soc Am*. 82:1933.

Ghitza O. 2012. On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation Spectrum. *Front Psychol*. 3:238.

Ghitza O, Greenberg S. 2009. On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*. 66:113–126.

Giraud A-L, Poeppel D. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*. 15:511–517.

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen MS. 2014. MNE software for processing MEG and EEG data. *NeuroImage*. 86(IS):446–460.

Green DM. 1960. Auditory detection of a noise signal. *J Acoust Soc Am*. 32:121.

Green DM. 1985. Temporal factors in psychoacoustics. Available at: http://link.springer.com/chapter/10.1007/978-3-642-70622-6_8.

Green DM, Swets JA. 1966. *Signal detection theory and psychophysics*. New York: Wiley.

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S. 2013. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol*. 11:e1001752.

Haegens S, Händel BF, Jensen O. 2011. Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *J Neurosci*. 31:5197–5204.

Henry MJ, Herrmann B, Obleser J. 2014. Entrained neural oscillations in multiple frequency bands comodulate behavior. *Proc Natl Acad Sci U S A*. 111:14935–14940.

Henry MJ, Obleser J. 2012. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc Natl Acad Sci U S A*. 109:20095–20100.

Herrmann B, Henry MJ, Grigutsch M, Obleser J. 2013. Oscillatory phase dynamics in neural entrainment underpin illusory percepts of time. *J Neurosci*. 33:15799–15809.

Jeffress LA. 1964. Stimulus-oriented approach to detection. *J Acoust Soc Am*. 36:766.

Jeffress LA. 1968. Mathematical and electrical models of auditory detection. *J Acoust Soc Am*. 44:187–203.

Johnson JS, Yin P, O'Connor KN, Sutter ML. 2012. Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis. *J Neurophysiol*. 107:3325–3341.

Kayser C, Ince RAA, Panzeri S. 2012. Analysis of slow (theta) oscillations as a potential temporal reference frame for information coding in sensory cortices. *PLoS Comput Biol*. 8:e1002717.

Kayser SJ, Ince RAA, Gross J, Kayser C. 2015. Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *J Neurosci*. 35:14691–14701.

Kerlin JR, Shahin AJ, Miller LM. 2010. Attentional gain control of ongoing cortical speech representations in a "cocktail party.". *J Neurosci*. 30:620–628.

Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ. 1999. Measuring phase synchrony in brain signals. *Hum Brain Mapp*. 8:194–208.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*. 320:110–113.

Lakatos P, Musacchia G, O'Connel MN, Falchier AY, Javitt DC, Schroeder CE. 2013. The spectrotemporal filter mechanism of auditory selective attention. *Neuron*. 77:750–761.

Lewicki MS. 2002. Efficient coding of natural sounds. *Nat Neurosci*. 5:356–363.

Luo H, Poeppel D. 2007. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*. 54:1001–1010.

Luo H, Poeppel D. 2012. Cortical oscillations in auditory perception and speech: evidence for two temporal windows in human auditory cortex. *Front Psychol*. 3:170.

Macmillan NA, Creelman CD. 2004. *Detection theory: a user's guide*. New York (NY): Taylor & Francis.

Moore BCJ. 1988. The shape of the ear's temporal window. *J Acoust Soc Am*. 83:1102.

Morillon B, Liégeois-Chauvel C, Arnal LH, Bénar C-G, Giraud A-L. 2012. Asymmetric function of theta and gamma activity in syllable processing: an intra-cortical study. *Front Psychol*. 3:248.

Narayan R, Graña G, Sen K. 2006. Distinct time scales in cortical discrimination of natural sounds in songbirds. *J Neurophysiol*. 96:252–258.

Nelken I, Rotman Y, Yosef OB. 1999. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*. 397:154–157.

Ng BSW, Logothetis NK, Kayser C. 2013. EEG phase patterns reflect the selectivity of neural firing. *Cereb Cortex*. 23:389–398.

Ng BSW, Schroeder T, Kayser C. 2012. A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci*. 32:12268–12276.

Obleser J, Weisz N. 2012. Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex*. 22:2466–2477.

Obleser J, Wöstmann M, Hellbernd N, Wilsch A, Maess B. 2012. Adverse listening conditions and memory load drive a common α oscillatory network. *J Neurosci*. 32:12376–12383.

Oldfield RC. 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 9:97–113.

Olshausen BA, Field DJ. 2004. Sparse coding of sensory inputs. *Curr Opin Neurobiol*. 14:481–487. Available at:. http://www.sciencedirect.com/science/article/pii/S0959438804001035/pdfft?md5=6db9d8d2058648631f63489ba83ed1f3&pid=1-s2.0-S0959438804001035-main.pdf.

Oostenveld R, Fries P, Maris E, Schoffelen J-M. 2011. Fieldtrip: open source software for advanced analysis of MEG, EEG,

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025

and invasive electrophysiological data. *Comput Intell Neurosci*. 2011:1–9.

Palva S, Palva JM, Shtyrov Y, Kujala T, Ilmoniemi RJ, Kaila K, Näätänen R. 2002. Distinct gamma-band evoked responses to speech and non-speech sounds in humans. *J Neurosci*. 22:RC211–RC211.

Panzeri S, Brunel N, Logothetis NK, Kayser C. 2010. Sensory neural codes using multiplexed temporal scales. *Trends Neurosci*. 33:111–120.

Peelle JE, Gross J, Davis MH. 2013. Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb Cortex*. 23:1378–1387.

Peña M, Melloni L. 2011. Brain oscillations during spoken sentence processing. *J Cogn Neurosci*. 24:1149–1164.

Plomp R, Bouman MA. 1959. Relation between hearing threshold and duration for tone pulses. *J Acoust Soc Am*. 31:749.

Poeppel D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time.". *Speech Comm*. 41:245–255.

Prins N, Kingdom FAA. 2009. Palamedes: Matlab routines for analyzing psychophysical data. http://www.palamedestoolbox.org.

Riecke L, Sack AT, Schroeder CE. 2015. Endogenous delta/theta sound-brain phase entrainment accelerates the buildup of auditory streaming. *Curr Biol*. 25:3196–3201.

Roberts TP, Ferrari P, Stufflebeam SM, Poeppel D. 2000. Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. *J Clin Neurophysiol*. 17:114–129.

Rosen S. 1992. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond Ser B Biol Sci*. 336:367–373.

Ross B, Miyazaki T, Thompson J, Jamali S, Fujioka T. 2014. Human cortical responses to slow and fast binaural beats reveal multiple mechanisms of binaural hearing. *J Neurophysiol*. 112:1871–1884.

Shahin AJ, Picton TW, Miller LM. 2009. Brain oscillations during semantic evaluation of speech. *Brain Cogn*. 70:259–266.

Singh NC, Theunissen FE. 2003. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am*. 114:3394–3411.

Smith EC, Lewicki MS. 2006. Efficient auditory coding. *Nature*. 439:978–982.

Strauß A, Wöstmann M, Obleser J. 2014. Cortical alpha oscillations as a tool for auditory selective inhibition. *Front Hum Neurosci*. 8:350.

Studebaker GA. 1985. A rationalized arcsine transform. *J Speech Lang Hear R*. 28:455–462.

Teng X, Tian X, Doelling K, Poeppel D. 2018. Theta band oscillations reflect more than entrainment: behavioral and neural evidence demonstrates an active chunking process. *Eur\ignorespacesJ\ignorespacesNeurosci*. 48:2770–2782.

Teng X, Tian X, Poeppel D. 2016. Testing multi-scale processing in the auditory system. *Sci Rep*. 6:34390.

Teng X, Tian X, Rowland J, Poeppel D. 2017. Concurrent temporal channels for auditory processing: oscillatory neural entrainment reveals segregation of function at different scales. *PLoS Biol*. 15:e2000812.

van Dijk H, Schoffelen J-M, Oostenveld R, Jensen O. 2008. Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci*. 28:1816–1823.

VanRullen R. 2006. The continuous wagon wheel illusion is associated with changes in electroencephalogram power at ~13 Hz. *J Neurosci*. 26:502–507.

VanRullen R, Koch C. 2003. Is perception discrete or continuous? *Trends Cogn Sci*. 7:207–213.

VanRullen R, Zoefel B, Ilhan BA. 2014. On the cyclic nature of perception in vision versus audition. *Philos Trans R Soc Lond B Biol Sci*. 369:20130214–20130214.

Viemeister NF. 1979. Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am*. 66:1364–1380.

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ. 2012. Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *J Neurophysiol*. 107:2033–2041.

Weisz N, Hartmann T, Müller N, Lorenz I, Obleser J. 2011. Alpha rhythms in audition: cognitive and clinical perspectives. *Front Psychol*. 2:73.

Wilsch A, Obleser J. 2016. What works in auditory working memory? A neural oscillations perspective. *Brain Res*. 1640:193–207. Available at:. http://dx.doi.org/10.1016/j.brainres.2015.10.054.

Wöstmann M, Herrmann B, Wilsch A, Obleser J. 2015. Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *J Neurosci*. 35:1458–1467.

Yekutieli D, Benjamini Y. 1999. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *J Stat Plan Infer*. 82:171–196.

Zoefel B, VanRullen R. 2015. Selective perceptual phase entrainment to speech rhythm in the absence of spectral energy fluctuations. *J Neurosci*. 35:1954–1964.

Zwislocki JJ. 1960. Theory of temporal auditory summation. *J Acoust Soc Am*. 32:1046.

Zwislocki JJ. 1969. Temporal summation of loudness: an analysis. *J Acoust Soc Am*. 46:431.

Downloaded from https://academic.oup.com/cercor/article/30/4/2600/5637582 by Head Serials and Electronic Resources Dept user on 28 May 2025