Report

Neural Response Phase Tracks How Listeners Learn New Acoustic Representations

Huan Luo,[1,4,*] Xing Tian,[2,4] Kun Song,[1] Ke Zhou,[1]
and David Poeppel[2,3]
[1]State Key Laboratory of Brain and Cognitive Science, Institute
of Biophysics, Chinese Academy of Sciences, Beijing 100101,
China
[2]Department of Psychology
[3]Center for Neural Science
New York University, New York, NY 10003, USA

Summary

Humans are remarkable at rapidly learning regularities
through experience from a dynamic environment [1, 2]. For
example, long-lasting memories are formed even for audi-
tory noise patterns after short, repeated exposure in an un-
supervised manner [3, 4]. Although animal neurophysiolog-
ical [5–10] and human studies [11–16] demonstrate adaptive
cortical plasticity after sensory learning and memory forma-
tion, the mechanisms by which the auditory system extracts
and encodes holistic patterns from random noise, which
contains neither semantic labels nor prominent acoustic fea-
tures to facilitate encoding, remains unknown. Here we com-
bined magnetoencephalography (MEG) with psychophysics
to address the issue. We demonstrate that the establishment
of a reliable neuronal phase pattern in low-frequency (3–
8 Hz) auditory cortical responses mirrors the noise memory
formation process. Specifically, with repeated exposure,
originally novel noise patterns are memorized, as reflected
in behavior, and gradually produce robust phase responses
in auditory cortex. Moreover, different memorized noises
elicit distinguishable phase responses, suggesting their
specificity to noise structure. The results indicate that the
gradual establishment of low-frequency oscillatory phase
patterns in auditory neuronal responses mediates the im-
plicit learning process by which originally undifferentiated
noises become new auditory objects.

Results

Memory Formation: Psychophysics

We employed an unsupervised noise memory paradigm,
adapted from a previous experiment [3], to study the dynamic
neural correlates of memory formation for white noise. As
illustrated in Figure 1A, listeners were either presented with a
noise sample generated by concatenating three identical
0.5 s noise segments (RN) or a 1.5 s running noise (N) and
were asked to determine the noise type (RN or N) by pressing
corresponding keys (Yes or No). Critically, one particular ex-
emplar of the RN sounds reoccurred, interspersed throughout
each experimental block (RefRN; red divided rectangle).
Similarly, one particular N sample was fixed and presented
repeatedly (RefN; red rectangle).

Figures 1B and 1C illustrate the mean behavioral perfor-
mance for 13 normal-hearing subjects. Consistent with the

previous provocative findings [3], RefRN showed higher
mean d′ scores (Wilcoxon signed rank test, p < 0.01; Figure 1B)
and shorter reaction times (Wilcoxon signed rank test, p < 0.01;
Figure 1C) than RN, suggesting that repeated exposure
successfully induced memory formation. Note that the learn-
ing was implicit and unsupervised, as subjects were only
required to judge the sound type (RN or N) and were not
made aware of the repeated exposure to RefRN and RefN in
each block. Furthermore, the sensitivity and reaction time as
a function of presentation order throughout a block revealed
a gradually developing course (Figures 1B and 1C, middle
panel), demonstrating that RefRN and RN were not initially
different and were only distinguishable after repeated expo-
sure to RefRN. On the other hand, RefN, although also
presented repeatedly (like RefRN), did not show memory for-
mation behavior (Figures 1B and 1C, right panel) in both
d′ (Wilcoxon signed rank test, p = 0.28) and reaction time
(Wilcoxon signed rank test, p = 0.27). A possible reason for
these differences in memory formation behavior between
RefRN and RefN is the decreased learning performance with
long sample duration [3].

Reoccurring Noise Elicits Robust Phase Pattern in Auditory Cortical Responses

We recorded magnetoencephalography (MEG) signals from
13 participants listening to the 1.5 s duration noise stimuli to
assess the neural correlates of noise memory development.
We first examined the canonical evoked response pattern in
auditory channels (independently defined by a tone localizer
in each subject separately), and as illustrated in Figure 2A
(across-channel root-mean-square [rms] waveforms), there
was no compelling difference for either RefRN/RN or RefN/N
pairs. Next, we performed a spectrotemporal analysis to
examine the induced power and phase responses as a
function of frequency and time. As shown in Figure 2B, we
did not observe noteworthy differences in induced power
responses for either RefRN/RN or RefN/N pairs. We did, how-
ever, find robust effects in the phase results. As illustrated in
Figure 2C (left), from ∼0.5 s after noise onset until appro-
ximately sound offset, RefRN produced stronger intertrial
phase coherence (ITC) in low frequency (2–8 Hz) responses
compared to RN (paired t test, p < 0.001), whereas neither
rms nor power analyses showed differences in this range
(paired t test, rms: p = 0.98; Power: p = 0.48). Interestingly,
RefN showed a similar trend of ITC advantage over N (Fig-
ure 2C, right) during this time-frequency period, although in
a weaker, noisier manner (paired t test, p = 0.03), paralleling
subjects' poor memory performance. These results thus sug-
gest a possible close link between low-frequency phase reli-
ability during this time-frequency range and noise learning,
leading us to focus on ITC within this "phase ROI" (0.5–1.5 s,
3–8 Hz, dotted box in left panel of Figure 2C) in subsequent
analyses. We did not incorporate the 2 Hz component into
the phase ROI to avoid the direct entrainment effect caused
by the repetitions in RefRN and RN sounds with segment
lengths of 0.5 s. Importantly, further analysis showed no
correlation in the ITC between 3–8 Hz and 2 Hz (Pearson's
correlation, n = 13, p = 0.12).

[4]These authors contributed equally to this work
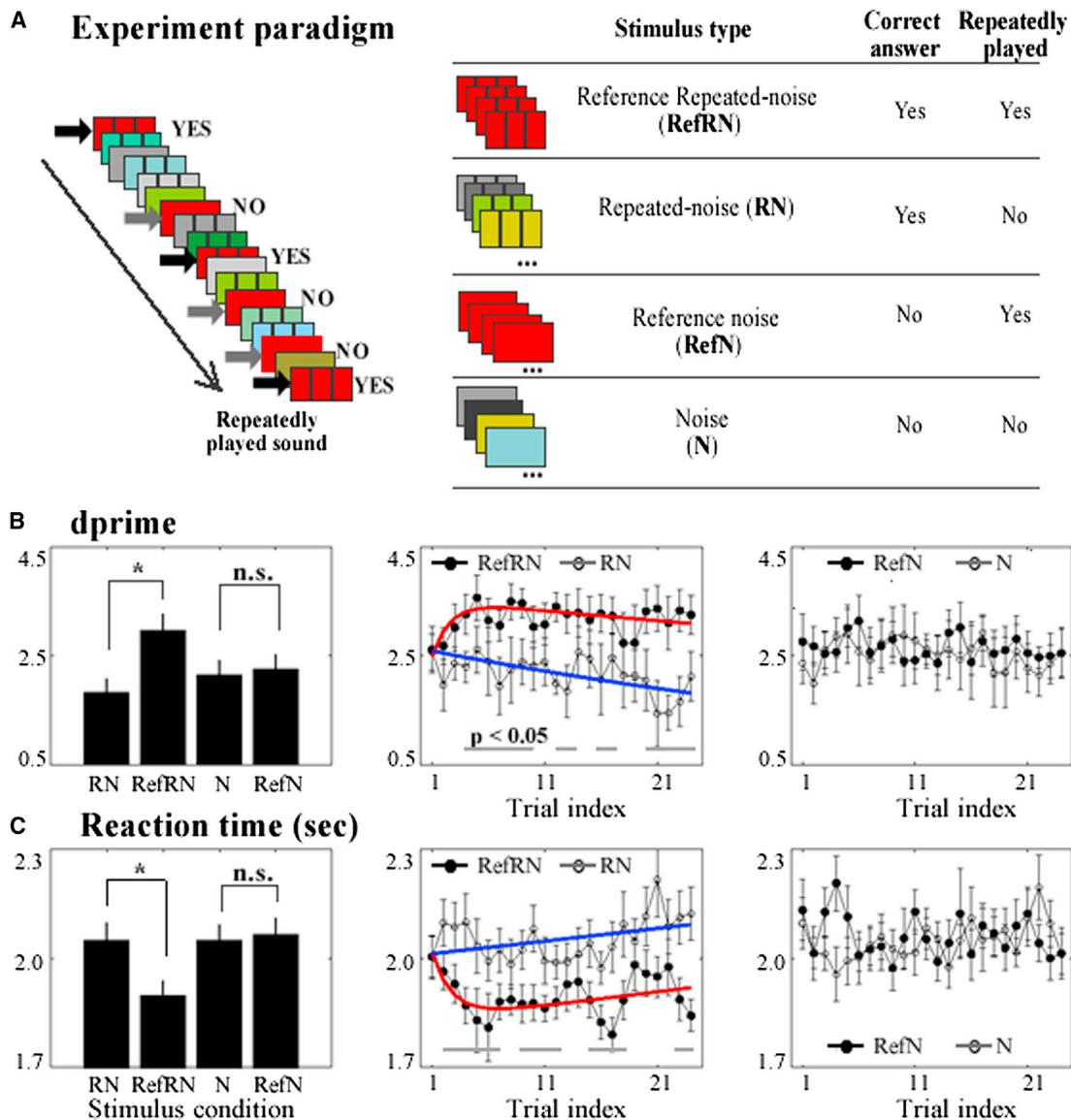*Correspondence: luohuan@gmail.com

Figure 1. Experimental Paradigm and Behavioral Results

(A) In each trial (rectangle) throughout each experiment block, subjects were asked to detect whether the 1.5 s noise sound contained repetitions (Yes or No). RN sounds were generated by seamlessly concatenating three identical 0.5 s noise segments and so contained repetitions (divided rectangles; correct answer: Yes). N sounds were 1.5 s running noise and so did not contain repetition (nondivided rectangle; correct answer: No). Different colors represent distinct noise exemplars. Within a block, one particular RN sound (RefRN, red divided rectangle and black arrows; correct answer: Yes) and one particular N sound (RefN, red nondivided rectangle and gray arrows; correct answer: No) were presented throughout a block repeatedly and identically. Accordingly, there were four stimulus conditions (RefRN, RN, RefN, and N).
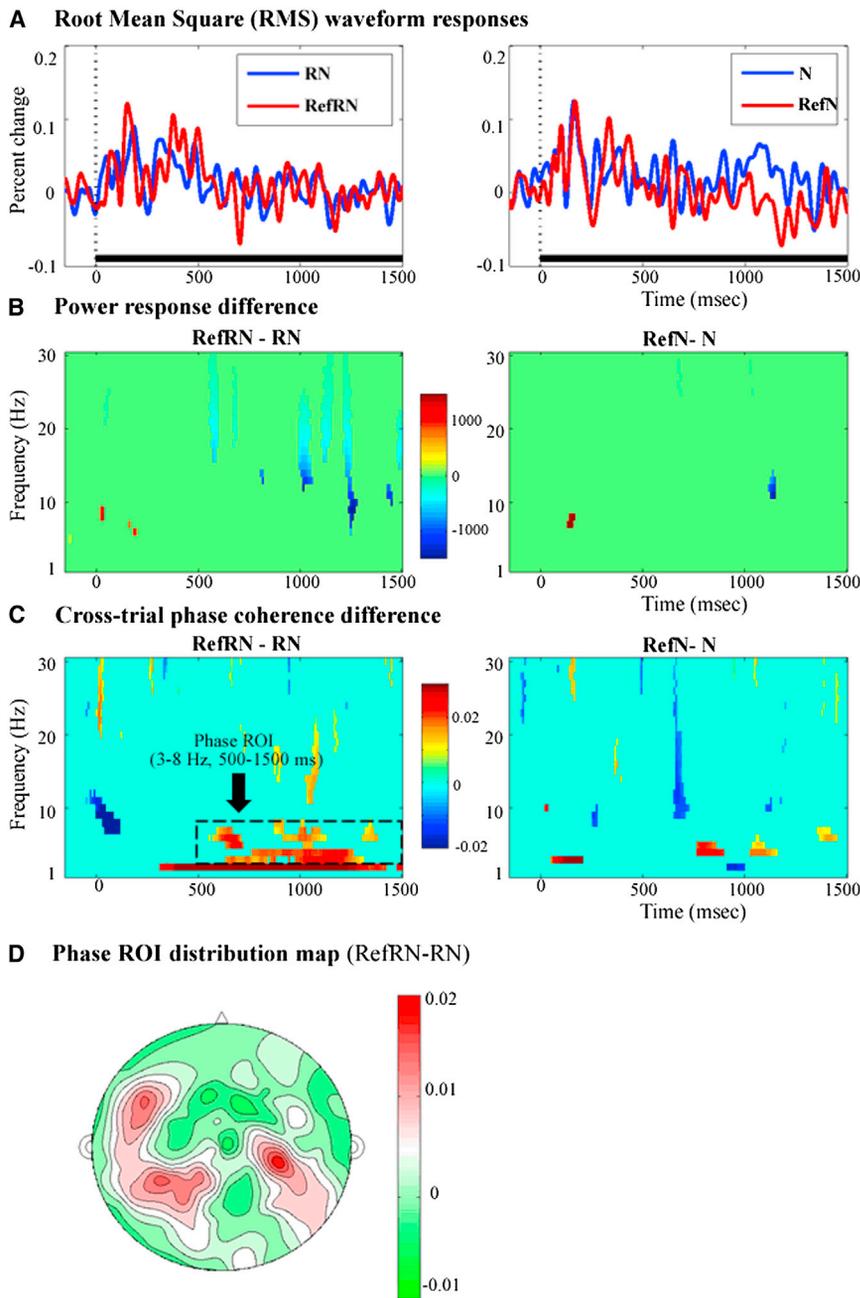
(B) Discrimination sensitivity (d') for RefRN, RN, RefN, and N. Overall performance (left, mean ± SEM, n = 13) and sensitivity as a function of trial order throughout experiment block (middle and right, mean ± SEM, n = 13). The solid lines show the exponential fitting for RefRN (red) and RN (blue). Black horizontal lines indicate significant differences between conditions (paired t test, n = 13, $p < 0.05$).

(C) Reaction time (RT) for RefRN, RN, RefN, and N (left, mean ± SEM, n = 13) and RT as a function of trial order (middle and right, mean ± SEM, n = 13).

We further examined the cortical spatial distribution of this 3–8 Hz phase pattern reliability by calculating the ITC difference between RefRN and RN, in the phase ROI, in each of the 157 MEG channels separately. As shown in Figure 2D (see also Figure S1, available online), a typical auditory cortex spatial distribution was observed, similar to that obtained from the independent tone localizer (significant spatial similarity, $p < 0.01$), further suggesting that RefRN elicited a robust low-frequency phase pattern in auditory cortex.

3–8 Hz Phase Reliability Buildup

We focused on the 3–8 Hz auditory cortical responses, by band-pass filtering the MEG responses of 20 auditory channels separately and then extracting the corresponding power and phase temporal patterns. As illustrated in Figure 3A, all four types of stimuli elicited a similar power pattern profile, with an initial onset response around 200 ms after sound onset and a subsequent sustained plateau throughout the noise presentation. Figure 3B summarizes the time course of inter-trial phase coherence results. All four stimuli elicited strong

A   Root Mean Square (RMS) waveform responses

B   Power response difference

C   Cross-trial phase coherence difference

D   Phase ROI distribution map (RefRN-RN)

Figure 2. Time-Frequency Responses and the Phase Coherence Difference Distribution Map

(A) Grand average (n = 13) evoked response (root-mean-square waveform responses across 20 auditory channels) for RefRN/RN pair (left) and RefN/N pair (right).

(B) Grand average time-frequency plots (n = 13, thresholded by paired t test, p ≤ 0.05, n = 13) for power difference (left: RefRN-RN; right: RefN-N) during 1.5 s stimulus presentation (black solid bar in A), with time 0 corresponding to the sound onset.

(C) Same as (B), but for intertrial phase coherence (ITC) difference. Dotted box indicates the time-frequency phase ROI (0.5–1.5 s after sound onset, 3–8 Hz) range that were focused on for further analysis.

(D) Grand average (n = 13) distribution map for mean ITC difference (RefRN-RN) in phase ROI (dotted box in C).

differ after RefRN reoccurred several times (e.g., after the seventh presentation). These results demonstrate that as noise occurred repetitively and was gradually encoded in memory throughout an experimental block, it progressively initiated a reliable phase pattern response in auditory cortex. Moreover, because each trial group contained identical RefRN stimuli (e.g., trials 1–6 versus trials 20–25), the results also support the perceptual correlates of phase pattern reliability. On the other hand, although having a similar trend, the RefN-N pair did not show as significant an evolving phase reliability as RefRN-RN did (Figure 3C, right), consistent with the psychophysical results. Furthermore, other time-frequency ROIs did not show similar ITC buildup, as in the phase ROI (Figures S2A–S2D).

Attention, Task Difficulty, and Segment Duration Controls

We ran two control experiments to address potential confounding factors (attention, task difficulty, and stimulus repetition structures) that could in part

phase reset after sound onset. Notably, matched with previous spectrogram results (Figure 2C), RefRN induced stronger ITC around 0.5 s after sound onset (Figure 3B, left) than RN, and RefN showed a weak trend of stronger phase reliability than N during the same time range (Figure 3B, right), which was defined as the phase ROI (box in Figure 3B).

Given that behavioral results showed a gradual memory buildup throughout each experimental block (Figures 1B and 1C, middle panel), we assessed whether the neuronal phase pattern for RefRN was established in a similar gradual, incremental manner. We calculated the ITC within the phase ROI, in steps of six trials within each experiment block (e.g., trials 1–6, 2–7, 3–8, etc.) and observed a gradual ITC buildup (Figure 3C, left). Specifically, the ITC was initially not different between RefRN and RN (e.g., the first six trials) and began to

account for the observed RefRN-RN ITC results. In the first control experiment, four subjects listened to the original stimuli as well as two additional new signals containing near-threshold 30 Hz amplitude modulation (AM noise), and they were asked to detect the AM noise instead of repetitions within a sound. All the original four stimuli (RN, RefRN, N, RefN) were thus nontargets, yet similar RefRN-RN ITC differences were observed (Figure 3D). In the second control experiment (n = 4), we increased the task difficulty and changed the noise segment duration by employing 1.6 s noise stimuli (RN: two concatenated 0.8 s noise segment; N: 1.6 s running noise). As shown in Figure 3E, in spite of increased task difficulty (d′ for RN: ∼0.5; d′ for RefRN: ∼1.5) and different segment duration (0.8 s instead of 0.5 s), we observed similar phase effects in 3–8 Hz (see Figure S2E for time-frequency plots),
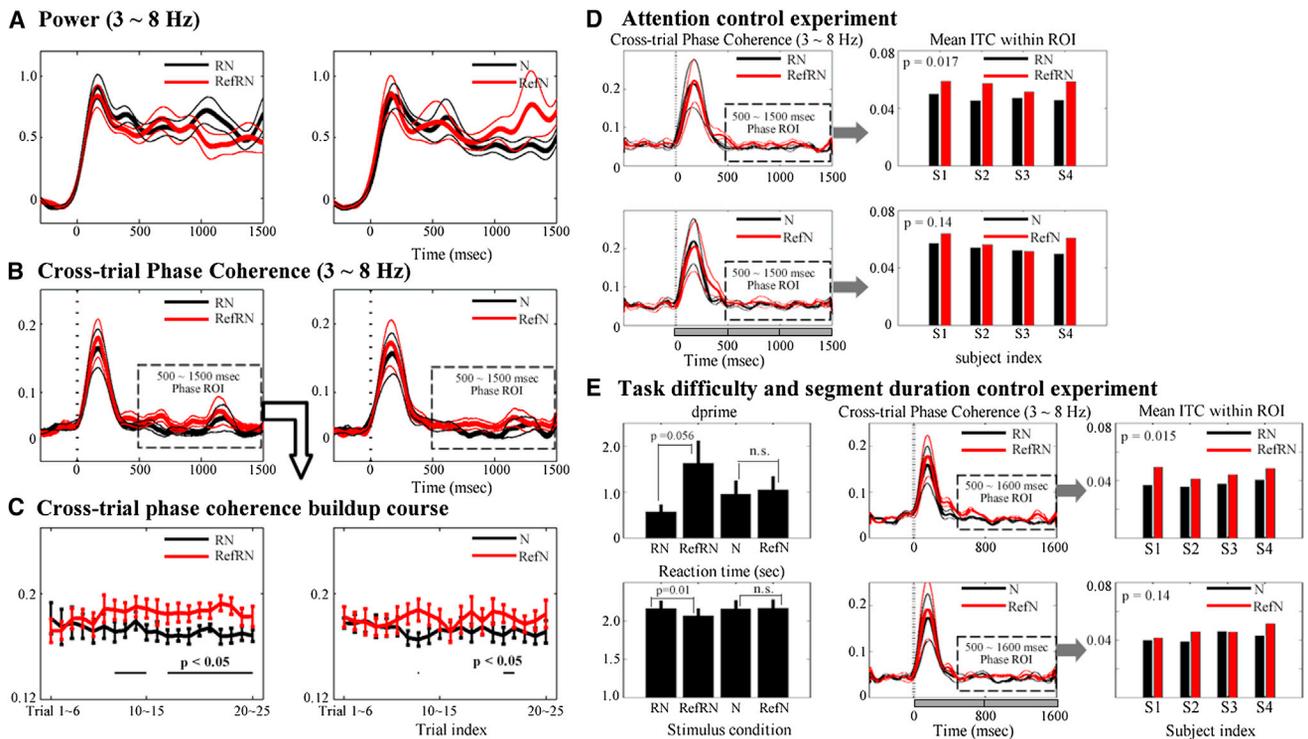
Figure 3. Three to Eight Herz Power and Phase Responses, Phase Reliability Buildup Course, and Performance in Two Control Experiments

(A) Grand average power responses (± SEM, n = 13) for RefRN/RN pair (left) and RefN/N pair (right) during stimulus presentation (0–1.5 s). Time 0 corresponds to stimulus onset.

(B) Same as (A) but for ITC responses. Dotted box indicates the phase ROI (0.5–1.5 s after sound onset).

(C) Grand average ITC in phase ROI (± SEM, n = 13) as a function of trial order within block for RefRN/RN pair (left) and RefN/N pair (right). Note that ITC was calculated across six consecutive trials (e.g., trials 1–6, 2–7, 20–25, etc.) instead of across 25 trials in (B) and thus had different ITC value range. Black horizontal lines indicate points showing significant differences between conditions (paired t test, p < 0.05).

(D) Attention control experiment. Left: Grand average 3–8 Hz ITC responses (± SEM, n = 4) in auditory channels for RefRN/RN pair (upper) and RefN/N pair (lower) during stimulus presentation (0–1.5 s). The dotted box indicates the phase ROI (0.5–1.5 s after sound onset). Right: Mean ITC in the phase ROI for each subject.

(E) Task difficulty and segment duration control experiment. Left: Discrimination sensitivity (d′) and reaction time (RT) (± SEM, n = 4). Middle: Grand average 3–8 Hz ITC responses for RefRN/RN pair (upper) and RefN/N pair (lower) during stimulus presentation (0–1.6 s). The dotted box indicates the phase ROI (0.5–1.6 s after sound onset). Right: Mean ITC in phase ROI for each subject. Note the stronger ITC for RefRN than RN in both control experiments.

arguing against the hypothesis that all effects are merely caused by phase entrainment.

Different Learned Noises Induce Diverse Phase Pattern Responses

Phase-mediated regularization in low-frequency neuronal oscillations has been previously found to be associated with attentional selection [17], and therefore it is natural to ask whether the observed robust phase pattern for RefRN was solely caused by the nonspecific top-down attentional modulation, given that RefRN would become familiar after implicit learning and in turn capture more attention. To test this possible interpretation, we examined whether the phase pattern responses were distinguishable among different RefRNs (all of them recurred, but in separate blocks) by performing a phase pattern discrimination analysis (discrimination score larger than zero indicates dissociated phase response) [18, 19]. The nonspecific attentional modulation hypothesis predicts similar, nondiscriminable phase patterns for different RefRN stimuli. As shown in Figure 4A, RefRN elicited phase responses that became distinguishable around 0.5 s after sound onset (left panel), approximately in the phase ROI range, whereas RefN did not produce differentiated

phase patterns (right panel), thus arguing against the nonspecific attention interpretation. Furthermore, we averaged the discrimination values within the phase ROI (0.5–1.5 s) to examine the overall discrimination ability. As shown in the left panel of Figure 4B, RefRN showed marginally significant discrimination scores (t test, p = 0.075; left bar). We further divided trials within one block into early and late trial groups (trials 1–12 and 13–24, respectively) to estimate the phase discrimination performances. Interestingly, we found that the phase responses for different RefRN sounds were established only in the later trial group (trials 13–24; t test, p = 0.03; right bar), but not in the early trial group (trials 1–12; t test, p = 0.046; middle bar) when memory has not been fully formed. On the other hand, RefN (Figure 4B, right panel) did not show discrimination ability in the total-trial (t test, p = 0.43), early-trial (t test, p = 0.38), and late-trial groups (t test, p = 0.78), suggesting that unlearned noises fail to elicit specific, distinguishable phase patterns. Thus, the diverse phase patterns for different RefRNs support the conjecture that the phase response was intimately associated with the sound structure rather than simply reflecting nonspecific attentional modulation and was shaped progressively as the memory was incrementally established.
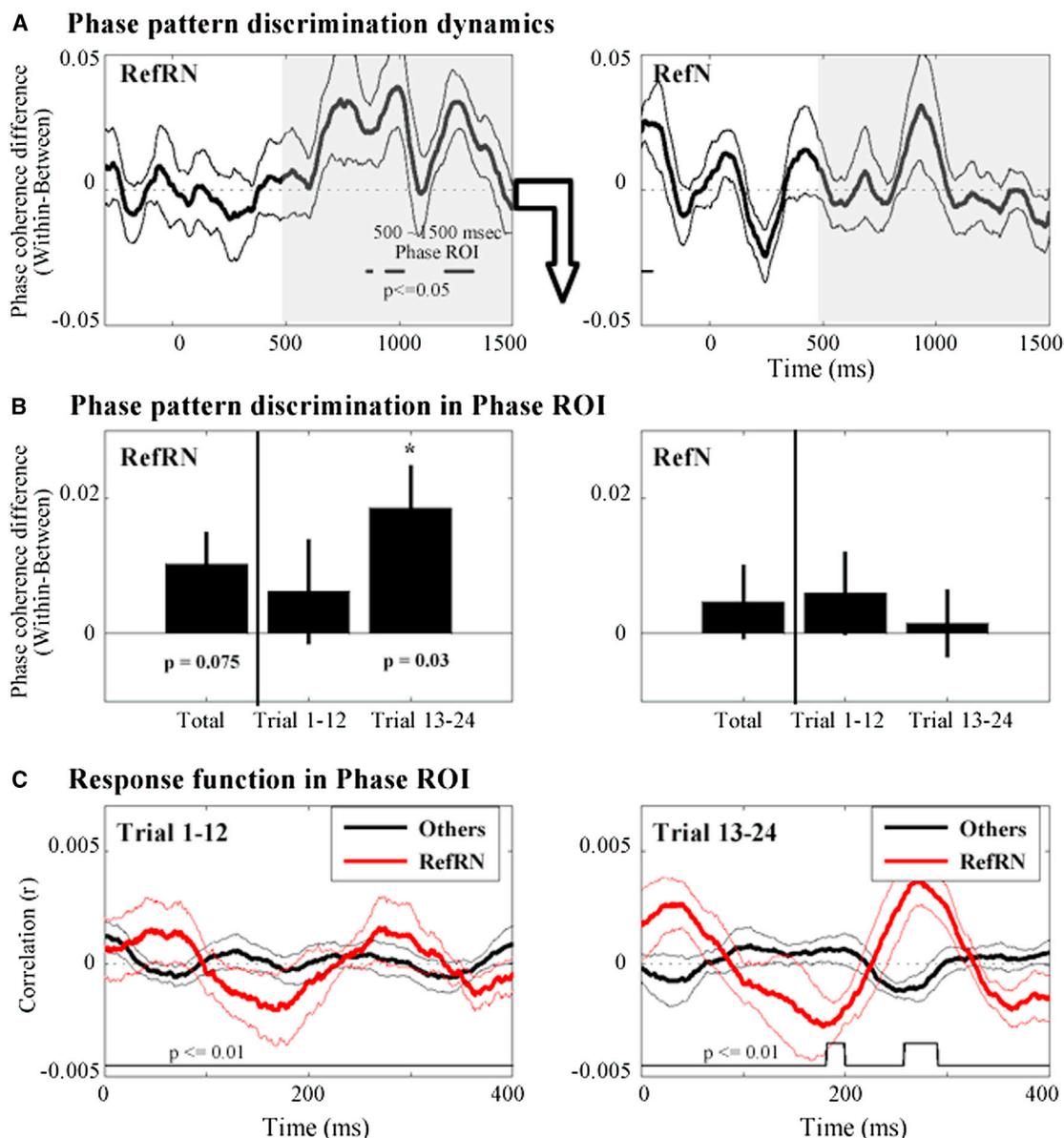
Figure 4. Dynamics of 3–8 Hz Phase Pattern Discrimination and Response Function

(A) Grand average phase pattern discrimination scores (± SEM, n = 6) for RefRN (left) and RefN (right). Time 0 indicates stimulus onset. Discrimination scores larger than 0 indicate distinguishable phase patterns for different RefRN sounds, excluding a general attentional modulation interpretation. Grey areas correspond to the phase ROI (0.5–1.5 s after sound onset). Black horizontal lines indicate significant nonzero values (paired t test, p < 0.05).

(B) Grand average phase pattern discrimination scores (± SEM, n = 6) averaged within phase ROI (gray area in A) for RefRN (left) and RefN (right), across all trials (trials 1–25) (total), the early-trial group (trials 1–12) and the late-trial group (trials 13–24). Stars indicate significant nonzero values (t test, p ≤ 0.05).

(C) Grand average response function ± SEM (n = 13) in early-trial (trials 1–12, left) and late-trial (trials 13–24, right) responses in phase ROI. Response functions of RN, N, and RefN were averaged (black, labeled as "others") and compared to that of RefRN (red). Black horizontal lines indicate time points showing significance differences between RefRN and others (paired t test, p < 0.01).

Response Function

Having established the *sensitivity* of the 3–8 Hz phase pattern to the noise learning process and documented its *specificity* to each RefRN, we next asked whether the repeated exposure induced the brain to become more sensitive to the subtle acoustic microstructures in the noise and in turn contributed to the formation of reliable phase responses. Such stimulus-response temporal relationships can be characterized by a "response function." Based on methods in recent visual studies [20, 21], we calculated the cross-correlation between the 3–8 Hz bandpass-filtered MEG responses within the phase ROI in each trial and the corresponding stimulus envelope, separately for the early-trial (trials 1–12) and late-trial (trials 13–24) groups, and for each stimulus condition, respectively (Figure S3). The response function for RefRN and the other sounds (averaged across the other three stimuli) did not differ in early trials when memory was not fully established (Figure 4C, left panel) and only became distinguishable during late trials (Figure 4C, right panel). Specifically, in the late-trial group, when RefRN was learned, the response function for

RefRN revealed a deflection starting around 200 ms compared to that for other sounds, whereas the early-trial group showed a similar but insignificant trend. Since the temporal waveform is determined by both phase and power pattern [22] and RefRN manifested strong phase reliability, the response function results presumably reflect contribution mainly from phase information. Hence, learning of a random noise after repeated exposure drove the auditory system to become more sensitive and phase-reset to subtle microstructures that were initially not obvious, in turn shaping a gradually emerging reliable phase pattern response.

Discussion

It is well established that sensory systems can take advantage of statistical regularities in the environment to achieve optimal performance through learning, memory, and attention [11, 23–25]. Converging neurophysiological recordings and noninvasive neuroimaging results suggest that sensory cortex undergoes rapid plasticity according to stimulus statistical context or task requirements [5, 26]. For example, the mismatch negativity response (MMN) in sensory cortex, which is elicited by oddball stimuli in the absence of an explicit task, reflects sensory memory and regularity extraction [15]. Moreover, auditory cortex has been found to mediate temporal expectation, another important form of regularity [9]; the statistical structure of sound sequences [10]; and perceptual organization [27]. Our results are therefore consistent with previous findings on the critical role of auditory cortex in regularity learning and structure formation. It is noteworthy that since the phase pattern response varies for different memorized noises, the results cannot be accounted for by nonspecific sustained responses to periodic noise [28] or awareness-related long-latency responses [16].

Recent neurophysiological studies demonstrate that cyclic cortical excitability states are reflected by oscillatory phase [25, 29]. Specifically, phase can act as a network-intrinsic reference frame for ongoing spike trains and therefore plays a critical role in sensory processing and attention, by temporally organizing continuous input into units of the appropriate temporal scale [17–19, 30–35]. Therefore, the observed gradually crystallizing specific phase pattern for learned sounds implicates the establishment of a temporal organization process in auditory processing, by regularizing spiking in time. Notably, all the previous studies [17–19, 30–35] employed stimuli containing explicit temporal structures (e.g., syllable onsets, lip movements, etc.), which may underpin robust phase resetting. On the other hand, unstructured noises did not drive successful phase tracking [18, 34], presumably due to the lack of segmentation cues. In the present study, we employed white noises that contain neither predictable nor prominent acoustic features—and yet demonstrated a gradual formation in phase response with learning, suggesting that random noises failed to elicit robust phase patterns initially when they were undifferentiated and elicited reliable phase responses as they developed into meaningful auditory objects.

An additional important aspect of our results lies in the fact that although RefNs did not form reliable memory traces as assessed by psychophysics, they showed a similar trend in phase reliability response development, although the established phase pattern was weaker and undifferentiated. In our view, the results thus indicate a rather general role of phase responses in the encoding of learned noises, which is independent of noise type (RN or N). It also suggests that if RefNs reoccurred for long-enough trials, they might also form specific distinguishable phase pattern responses.

Given the temporal nature of audition, an auditory object is often regarded as a stream that builds up over time (unlike a static image in vision) and therefore requires the neural representation to capture the ongoing temporal structure. It has been recently postulated that temporal coherence across various sound features, together with attention, plays a binding role in auditory object formation [36]. Interestingly, this view can be naturally linked to the phase results here, because phase coordination among neuronal populations encoding different acoustic features could assist their mutual communication by aligning their high excitability [37]. Our findings are consistent with a population-level temporal encoding mechanism—phase-mediated temporal coordination—that underlies auditory object representation. Moreover, we observed reliable phase tracking in the 3–8 Hz frequency range of neuronal oscillations, consistent with many auditory studies [18, 30, 32, 33, 35], whereas visual research has instead disclosed the critical role of high frequencies [37, 38], possibly due to differences between sensory modalities.

In summary, we report that as a frozen noise reoccurred and became differentiated from other noise sounds, it gradually initiated a specific robust phase pattern in auditory cortical responses. The results demonstrate the important role of oscillatory phase pattern in encoding of auditory objects and suggest a population-level neural representation for newly learned auditory objects, based on phase-mediated organization pattern in time, by analogy with spatial organization in vision.

Experimental Procedures

A brief description of experimental procedures is in Results. A complete description can be found in the Supplemental Experimental Procedures.

Supplemental Information

Supplemental Information includes three figures and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.04.031.

References

1. Saffran, J.R., Johnson, E.K., Aslin, R.N., and Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. Cognition 70, 27–52.
2. Turk-Browne, N.B., Jungé, J., and Scholl, B.J. (2005). The automaticity of visual statistical learning. J. Exp. Psychol. Gen. 134, 552–564.
3. Agus, T.R., Thorpe, S.J., and Pressnitzer, D. (2010). Rapid formation of robust auditory memories: insights from noise. Neuron 66, 610–618.
4. Seitz, A.R. (2010). Sensory learning: rapid extraction of meaning from noise. Curr. Biol. 20, R643–R644.

5. Buonomano, D.V., and Merzenich, M.M. (1998). Cortical plasticity: from synapses to maps. Annu. Rev. Neurosci. *21*, 149–186.

6. Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat. Neurosci. *6*, 1216–1223.

7. Bao, S., Chang, E.F., Woods, J., and Merzenich, M.M. (2004). Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. Nat. Neurosci. *7*, 974–981.

8. Schnupp, J.W., Hall, T.M., Kokelaar, R.F., and Ahmed, B. (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. J. Neurosci. *26*, 4785–4795.

9. Jaramillo, S., and Zador, A.M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. Nat. Neurosci. *14*, 246–251.

10. Yaron, A., Hershenhoren, I., and Nelken, I. (2012). Sensitivity to complex statistical regularities in rat auditory cortex. Neuron *76*, 603–615.

11. Harris, J.A., Petersen, R.S., and Diamond, M.E. (2001). The cortical distribution of sensory memories. Neuron *30*, 315–318.

12. Jäncke, L., Gaab, N., Wüstenberg, T., Scheich, H., and Heinze, H.J. (2001). Short-term functional plasticity in the human auditory cortex: an fMRI study. Brain Res. Cogn. Brain Res. *12*, 479–485.

13. Tremblay, K., Kraus, N., McGee, T., Ponton, C., and Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. Ear Hear. *22*, 79–90.

14. Pleger, B., Foerster, A.F., Ragert, P., Dinse, H.R., Schwenkreis, P., Malin, J.P., Nicolas, V., and Tegenthoff, M. (2003). Functional imaging of perceptual learning in human primary and secondary somatosensory cortex. Neuron *40*, 643–653.

15. Gottselig, J.M., Brandeis, D., Hofer-Tinguely, G., Borbely, A.A., and Achermann, P. (2004). Human central auditory plasticity associated with tone sequence learning. Learn Mem. *11*, 162–171.

16. Gutschalk, A., Micheyl, C., and Oxenham, A.J. (2008). Neural correlates of auditory perceptual awareness under informational masking. PLoS Biol. *6*, e138.

17. Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., and Schroeder, C.E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. Science *320*, 110–113.

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

19. Luo, H., Liu, Z., and Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. PLoS Biol. *8*, e1000445.

20. Lalor, E.C., Pearlmutter, B.A., Reilly, R.B., McDarby, G., and Foxe, J.J. (2006). The VESPA: a method for the rapid estimation of a visual evoked potential. Neuroimage *32*, 1549–1561.

21. VanRullen, R., and Macdonald, J.S. (2012). Perceptual echoes at 10 Hz in the human brain. Curr. Biol. *22*, 995–999.

22. Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E., and Sejnowski, T.J. (2002). Dynamic brain sources of visual evoked responses. Science *295*, 690–694.

23. Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. *8*, 457–464.

24. Seitz, A.R., and Dinse, H.R. (2007). A common framework for perceptual learning. Curr. Opin. Neurobiol. *17*, 148–153.

25. Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. *32*, 9–18.

26. Jääskeläinen, I.P., Ahveninen, J., Belliveau, J.W., Raij, T., and Sams, M. (2007). Short-term plasticity in auditory cognition. Trends Neurosci. *30*, 653–661.

27. Micheyl, C., Tian, B., Carlyon, R.P., and Rauschecker, J.P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron *48*, 139–148.

28. Keceli, S., Inui, K., Okamoto, H., Otsuru, N., and Kakigi, R. (2012). Auditory sustained field responses to periodic noise. BMC Neurosci. *13*, 7.

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

30. Kayser, C., Montemurro, M.A., Logothetis, N.K., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron *61*, 597–608.

31. Masquelier, T., Hugues, E., Deco, G., and Thorpe, S.J. (2009). Oscillations, phase-of-firing coding, and spike timing-dependent plasticity: an efficient learning scheme. J. Neurosci. *29*, 13484–13493.

32. Szymanski, F.D., Rabinowitz, N.C., Magri, C., Panzeri, S., and Schnupp, J.W. (2011). The laminar and temporal structure of stimulus information in the phase of field potentials of auditory cortex. J. Neurosci. *31*, 15787–15801.

33. Kayser, C., Ince, R.A., and 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.

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

35. Ng, B.S., Schroeder, T., and Kayser, C. (2012). A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. J. Neurosci. *32*, 12268–12276.

36. Shamma, S.A., Elhilali, M., and Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. Trends Neurosci. *34*, 114–123.

37. Panzeri, S., Brunel, N., Logothetis, N.K., and Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. Trends Neurosci. *33*, 111–120.

38. Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. Annu. Rev. Neurosci. *32*, 209–224.