Perceptual and Semantic Contributions to Repetition Priming of Environmental Sounds

Marzia De Lucia[1], Luca Cocchi[2,3], Roberto Martuzzi[4], Reto A. Meuli[4], Stephanie Clarke[5] and Micah M. Murray[1,4,5]

[1]Electroencephalography Brain Mapping Core, Center for Biomedical Imaging, Centre Hospitalier Universitaire Vaudois and University of Lausanne, 1011 Lausanne, Switzerland, [2]Institute of Psychology, University of Lausanne, 1005 Lausanne, Switzerland, [3]Melbourne Neuropsychiatry Centre, University of Melbourne, 3053 Melbourne, Australia, [4]Radiology Service and [5]Neuropsychology and Neurorehabilitation Service, Centre Hospitalier Universitaire Vaudois and University of Lausanne, 1011 Lausanne, Switzerland

Repetition of environmental sounds, like their visual counterparts, can facilitate behavior and modulate neural responses, exemplifying plasticity in how auditory objects are represented or accessed. It remains controversial whether such repetition priming/suppression involves solely plasticity based on acoustic features and/or also access to semantic features. To evaluate contributions of physical and semantic features in eliciting repetition-induced plasticity, the present functional magnetic resonance imaging (fMRI) study repeated either identical or different exemplars of the initially presented object; reasoning that identical exemplars share both physical and semantic features, whereas different exemplars share only semantic features. Participants performed a living/man-made categorization task while being scanned at 3T. Repeated stimuli of both types significantly facilitated reaction times versus initial presentations, demonstrating perceptual and semantic repetition priming. There was also repetition suppression of fMRI activity within overlapping temporal, premotor, and prefrontal regions of the auditory "what" pathway. Importantly, the magnitude of suppression effects was equivalent for both physically identical and semantically related exemplars. That the degree of repetition suppression was irrespective of whether or not both perceptual and semantic information was repeated is suggestive of a degree of acoustically independent semantic analysis in how object representations are maintained and retrieved.

Keywords: auditory, fMRI, object recognition, perceptual priming, semantic priming, what and where pathways

Introduction

The human auditory system quickly and accurately recognizes sounds of objects in the environment as well as speech even in noisy conditions (Bregman 1990). These processes are subject to learning and plasticity, such as through repeated stimulus exposure that in turn can produce experience-dependent modulations in neural activity (reviewed in Grill-Spector et al. 2006). We investigated the neurophysiologic mechanisms contributing to such learning and plasticity in an effort to better understand how auditory objects are represented and reaccessed.

Stimulus repetitions typically facilitate reaction times and accuracy rates. Such effects, commonly referred to as repetition priming, are observed across sensory modalities and constitute a form of implicit memory (Henson 2003; Schacter et al. 2004). Two classes of repetition priming have been described (Schacter and Buckner 1998). Perceptual priming is linked to the physical features of the stimulus, such that changes to these features across initial and repeated stimulus exposures reduces, if not eliminates altogether, the behavioral facilitation. Conceptual or semantic priming occurs despite such changes and is instead linked to the underlying referent (i.e., the object itself). While both classes of priming have been documented using visual and linguistic (both visual and acoustic) stimuli, it remains controversial as to whether semantic priming can be elicited with sounds of environmental objects (Stuart and Jones 1995; Chiu 2000).

Parallel controversies exist concerning the neurophysiologic basis of priming. Whether repetition suppression and/or repetition enhancement is linked to the behavioral phenomenon of priming may vary as a function of brain region, latency poststimulus, and stimulus materials/task (Schacter et al. 2004; Wig et al. 2005; Grill-Spector et al. 2006). With regard to auditory stimuli, neuroimaging investigations have almost exclusively utilized linguistic stimuli and have obtained priming-related effects within extrastriate (i.e., visual) and prefrontal cortices (Buckner et al. 2000; Badgaiyan et al. 2001). The predominant interpretation is that such extrastriate regions mediate priming irrespective of the sensory modality and also despite changes in the surface features (i.e., acoustics; Badgaiyan et al. 2001). The implication is that common regions and mechanisms are involved in both perceptual and semantic priming of auditory and visual stimuli (Schacter et al. 2004). More recently, it has been shown that auditory cortices of the temporal lobe are involved in perceptual priming of sounds of environmental objects (Bergerbest et al. 2004; Murray et al. 2008; also Ahveninen et al. 2006 for effects involving syllables), suggesting that priming sounds of environmental objects might instead recruit distinct networks from what has been previously observed with either linguistic auditory or visual object stimuli. Specifically, repetition-induced plasticity in representations of sounds of environmental objects would appear to recruit temporal lobe structures traditionally associated with auditory functions.

One unresolved issue, and the focus of the present study, is whether these representations are reflecting the acoustic and/or semantic features of the objects. An argument favoring plasticity in semantic features is the timing of effects on auditory evoked potentials (AEPs). Murray et al. (2008) observed repetition suppression of AEPs at ~160-ms poststimulus onset, which is ~80 ms after the initial semantic discrimination of objects. Plus, the localization of these repetition suppression effects was predominantly within brain regions of the left middle temporal cortices (for evidence from an adaptation paradigm, see also Altmann et al. 2007), whereas the earlier categorical effects predominantly modulated right middle temporal cortices. Consequently, Murray et al. (2008)

proposed that such repetition priming/suppression involves not only simply plasticity based on pure acoustic features but also (at least minimal) access to some semantic features. Others, focusing on immediate repetition (adaptation) paradigms, have proposed that modulations of the evoked magnetic fields over the 150- to 250-ms poststimulus period are linked to spectral, rather than either temporal or semantic features (Altmann et al. 2008). The present event-related functional magnetic resonance imaging (fMRI) study therefore directly assessed whether representations of environmental sounds are subject to plasticity in an abstract or solely exemplar-specific manner. Similar behavioral and neural (vis a vis fMRI effects) priming effects in response to repetition of either the same acoustic exemplar or a different exemplar of the same object would support the hypothesis that repetition-induced plasticity involves access to semantic representations.

Materials and Methods

Participants

Twenty healthy, right-handed individuals (8 females), aged 22–36 years, participated (mean ± standard deviation [SD] = 28.1 ± 4.2 years). None had a history of neurological or psychiatric illnesses, and all reported normal hearing. All participants provided informed consent to participate in the study, the procedures of which were approved by the Ethics Committee of the Faculty of Biology and Medicine at the University of Lausanne. One participant (26-year-old male) failed to exhibit robust activations to the auditory stimuli (vs. rest), and his behavioral and fMRI data were therefore excluded from group analyses. All data presented in this study are from the remaining 19 individuals.

Stimuli

Auditory stimuli were complex, meaningful sounds (16-bit stereo; 22 500 Hz digitization) of common and readily categorized environmental objects taken from a library of sounds used in our prior studies (c.f. Murray et al. 2006 for a full listing, including details on the psychometrics). From this original library of 120 sound files, 48 were included in the present study because the original library includes several exemplars of the same object. These 48 sound files constituted sounds of 32 distinct objects—that is, for 16 of the objects, there were 2 different exemplars. Half of the sounds represented sounds of living objects (animal vocalizations, nonverbal human vocalizations, etc.), and the other half represented sounds of man-made objects (musical instruments, vehicles, and household objects, etc.). Each sound was 500 ms in duration, which included an envelope of 50-ms decay time that was applied to the end of the sound file to minimize clicks at sound offset. All sounds were further normalized according to the root mean square of their amplitude. Note that unlike prior studies, we have used relatively short-duration sound stimuli (2 s in Bergerbest et al. 2004; 5 s in Chiu 2000; and 12 s in Stuart and Jones 1995). Short-duration stimuli likely minimize the extent of conceptual processing and by extension possible intersubject heterogeneity in the activated networks. In addition, physical attributes of auditory objects can be more readily controlled (c.f. Murray et al. 2006). In the case of the present study, acoustic analyses of the stimuli were done by statistically comparing the spectrograms (using Matlab's spectrogram function with no overlapping and zero padding), using a time-frequency bin width of ~5 ms and ~50 Hz (see also Aeschlimann et al. 2008; Knebel et al. 2008). Statistical contrasts entailed a series of nonparametric t-tests based on a bootstrapping procedure with 1000 iterations (Manly 1991). To partially correct for multiple contrasts and autocorrelation, a significant difference at a given time-frequency bin was only considered reliable if all 8 of its immediately adjacent bins also yielded P values ≤ 0.05 (i.e., a 3 × 3 bin threshold was applied). There were no statistically reliable differences at any time-frequency bin between the spectrograms from between the different subgroups of initial sound presentations (see Supplementary Fig. S1a). Consequently, these spectrograms were

pooled and subsequently compared with those from the repeated sounds. This was done separately for repetitions of the physically identical sounds and for repetitions of the semantically identically sounds (see Supplementary Fig. S1b and S1c, respectively). Once again, no statistically reliable differences were observed.

Stimuli were likewise analyzed in terms of their mean harmonics-to-noise ratio (HNR), which was calculated using PRAAT software (http://www.fon.hum.uva.nl/praat/). This is a method to quantify and compare dynamic acoustic properties (i.e., periodicity) of sounds (Lewis et al. 2005). The mean (±standard error of the mean [SEM]) HNR for initially presented sounds did not significantly differ ($P > 0.65$) between those that would be repeated identically (9.1 ± 1.5; range 2.8–22.6) and those that would be repeated semantically (10.4 ± 2.5; range –2.0 to 27.1). Consequently and as above, for this analysis, we pooled all of the sounds presented initially together (mean HNR ± SEM = 9.7 ± 1.4; range –2.0 to 27.1) and compared the HNR from this pooled group with that from the group of sounds repeated semantically (10.9 ± 2.3; range –3.0 to 24.9). There was no evidence of a reliable difference ($P > 0.68$).

In addition, we evaluated the extent to which the auditory objects that were repeated with a different exemplar were indeed acoustically different from their initial counterparts. This control is necessary for excluding the possibility that any semantic priming effect is (partially) driven by similar acoustic features between the initial and repeated sounds. We therefore contrasted the spectrograms of these 2 groups of sounds using a paired t-test based on a bootstrapping procedure with 1000 iterations (Manly 1991) as described above. To establish a basis of comparison, this analysis was also conducted contrasting the 2 subgroups of initial sounds—that is, those whose repetition was acoustically identical and those whose repetition was acoustically different. This paired t-test was conducted 10 times based on a different random order of these 2 groups of sounds. The aim was to have an estimation of the extent of the difference in the spectrograms between sounds that were not related semantically and to evaluate whether the "acoustic distance" between the initial and repeated semantic stimuli was comparable. This yields the number of significantly differing frequency bins as a function of time during the stimulus. These results are displayed in the upper panel of Supplementary Figure S2 for the 10 shuffled data sets (blue lines) as well as for the auditory objects that were repeated with a different exemplar versus their initial counterparts (red line). These results show that this last analysis falls within the range determined from the shuffled data sets, which refer to different objects. Thus, the spectrotemporal difference between different exemplars of the same object is comparable to that between sounds of different objects. We also tested whether temporal shifts of similar frequency features could explain the differences between exemplars of the same objects. We did this by artificially "shifting" in time (in 25-ms steps) the time-frequency representations of the initial with respect to the repeated sounds of the same objects. The results show that any acoustic differences due to shifts in time fall within the same range as those observed when comparing different objects.

Stimuli were presented to participants via MR-compatible piezoelectric headphones (Nordic NeuroLab, Bergen, Norway). Sound volume was adjusted to a level deemed comfortable by each participant from within the bore of the scanner prior to the start of the experiment, such that stimuli were clearly audible and distinguishable. Task performance (see Results) provides a post hoc indication that this was indeed the case.

Procedure and Task

Sounds were presented one at a time while participants performed a living versus man-made forced-choice discrimination. Each participant completed a single block of 64 trials, which lasted approximately 6.5 min. The sound of a given object was presented twice during the block of trials, and the second presentation was either the physically identical sound or was a physically different exemplar of the same object. We hereafter refer to the initial presentation of sounds with the label INIT throughout the remainder of the text. Because physically identical sound repetitions are also semantically identical, we hereafter refer to this condition as "physical semantic repetition," using the label PSR for simplicity. Likewise, we refer to semantic repetitions with the label SR. The paradigm thus followed a 1 × 3 within-subjects design,

with the factor of presentation type (INIT, PSR, and SR). The mean (±SD) number of intervening trials between initial and repeated presentations of the physically identical sound was 5.2 ± 4.2 trials and between initial and repeated presentations of physically different exemplars of the same object was 4.0 ± 3.1 trials. These values did not significantly differ ($P > 0.35$). In addition, the mean trial number (i.e., the position within the block of trials) when stimuli were initially presented did not significantly differ between those stimuli that were subsequently repeated in a physically identical manner and those that were repeated in a semantically identical manner ($P > 0.10$). The interstimulus interval ranged from 2500 to 6000 ms in steps of 500 ms and was equally distributed for all experimental conditions. E-prime (Psychology Software Tools, Inc., Pittsburgh, PA) was used to control stimulus delivery and to record behavioral responses; the latter of which were acquired from an MR-compatible keypad (Photon Control Inc., Burnaby, BC, Canada).

Magnetic Resonance Imaging

Functional MRI data were acquired using an event-related design on a 3.0-T Siemens Trio system equipped with a 12-channel head coil. Blood oxygen level–dependent (BOLD) signals were obtained with a single-shot gradient echo planar imaging (EPI) sequence (time repetition [TR] = 2 s, time echo [TE] = 30 ms, field of view [FOV] = 224 mm, flip angle = 90°, matrix size 64 × 64). Each volume was comprised of 36 slices (slice thickness 3 mm; gap = 0.3 mm) covering the entire cerebral hemispheres and acquired in descending order (i.e., first slice at the top of the head). To provide precise structural and anatomical localization of brain activity, a sagittal T_1-weighted 3D gradient echo sequence was acquired for each subject (160 contiguous sagittal slices, slice thickness 1 mm, matrix size 256 × 256, TR = 1.48 s, TE = 2.63 ms, FOV = 256 mm, flip angle = 15°).

Data Analysis

Activation maps were obtained using SPM5 software (Wellcome Department of Cognitive Neurology, London, UK). Spatial realignment to the first volume and temporal realignment were applied to all the functional volumes for each subject. Functional images were then normalized to a standard space, as defined by the Montreal Neurological Institute EPI template, resampled to a voxel size of 3 × 3 × 3 mm³, and finally smoothed with an isotropic Gaussian kernel (full width half maximum = 6 mm). Structural images were first coregistered to the functional data and then normalized to the standard space applying the same transformation used for the functional volumes. The statistical analysis was performed with the general linear model (Friston et al. 1994), using the canonical hemodynamic response as basic function and its temporal derivative, as defined in SPM5. Demeaned reaction time (based on the mean across conditions for a given participant) was included as a parametric modulator for each subject and condition in order to account for nonspecific effects (e.g., attention, decision making, response selection, error monitoring, etc.).

Activation maps at the group level were inferred by means of second-level statistics, according to the random effects theory (Holmes and Friston 1998). Analyses were conducted to determine changes related to stimulation condition (INIT, PSR, and SR) based on a 1-way repeated measures analysis of variance (ANOVA) with the within-subjects factor of presentation type (voxel-level threshold at $P < 0.001$, family-wise error [FWE] corrected; 5 voxel spatial extent threshold). Post hoc analyses were limited to those clusters showing a significant main effect ANOVA and were obtained by means of paired t-tests between conditions (voxel-level threshold at $P < 0.001$, FWE corrected; $k_E = 5$ voxels).

Results

Behavioral Results

Participants readily completed the task, and accuracy rates were over 90% for each experimental condition (mean ± SEM for INIT = 91.6 ± 1.5%; PSR = 91.8 ± 1.4%; SR = 91.4 ± 2.0%). A 1 × 3 within-subjects repeated measures ANOVA was performed on accuracy rates and failed to reveal a main effect ($F_{2,17} = 0.47$; $P > 0.62$), providing no evidence of repetition priming on living versus man-made categorical discrimination accuracy.

A similar analysis of reaction times (Fig. 1a), by contrast, did identify a significant main effect of presentation type ($F_{2,17} = 12.960$; $P < 0.001$). Post hoc contrasts (paired t-tests) confirmed that reaction times were significantly facilitated for PSR ($t_{18} = 3.21$; $P < 0.005$) and SR ($t_{18} = 4.38$; $P < 0.0004$) versus INIT, though no difference between repetition types was observed ($t_{18} = 0.96$; $P > 0.35$). That is, we observed both perceptual and semantic priming effects on mean reaction times.
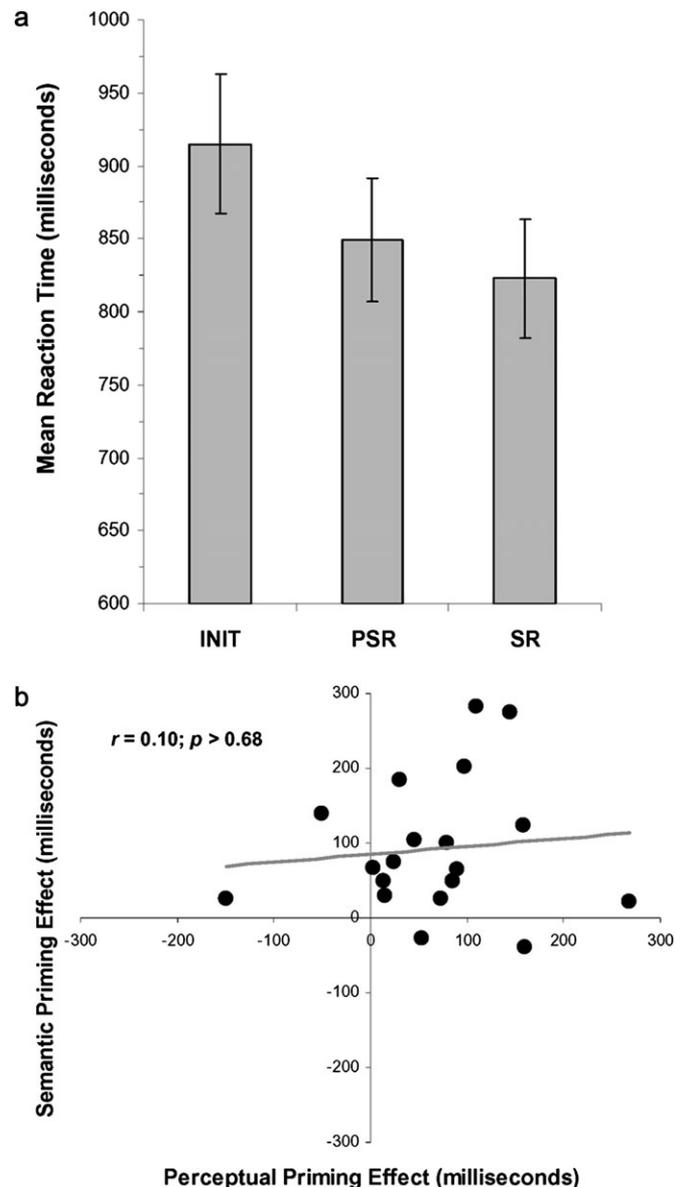

Figure 1. Behavioral results. (a) The top panel illustrates group-averaged ($N = 19$; SEM indicated) reaction time data as a function of presentation type (INIT, PSR, and SR). (b) The bottom panel plots the reaction time difference between initial and repeated sound presentations when the exemplars were semantically identical (vertical axis; SR condition) versus the reaction time difference between initial and repeated sound presentations when the exemplars were physically and semantically identical (horizontal axis; PSR condition), which we refer to as semantic and perceptual priming effects, respectively. There was no evidence for a linear relationship between these priming effects (Pearson correlation coefficient indicated).

We then calculated the reaction time difference between initial and repeated presentations for each priming variety and for each participant. This allowed us to evaluate whether the magnitude of perceptual and semantic priming effects was linearly correlated. The mean (±SEM) perceptual priming effect across participants was 66 ± 20 ms, and the mean semantic priming effect across individuals was 92 ± 21 ms. As indicated by the above ANOVA and post hoc contrasts, these values did not significantly differ (t_{18} = 0.96; P > 0.35). Moreover, analyses of single-subject data sets revealed that only 3 of the 19 participants exhibited significant reaction time (RT) differences between PSR and SR conditions, though there was no consistent pattern in terms of directionality. Likewise, there was no evidence for a significant correlation between perceptual and semantic priming effects (r_{17} = 0.10; P > 0.68; Fig. 1b). In other words, there was no evidence that the magnitude of perceptual priming was indicative of the magnitude of semantic priming at a behavioral level. It should be noted that had a significant correlation been observed, it could have been largely (if not wholly) explained by the fact that the same data (i.e., the INIT condition) contributes to both priming effects. The fact that we failed to observe a reliable correlation can thus be considered robust.

fMRI Results

Figure 2 displays the activation patterns for each condition versus baseline (P < 0.001, FWE corrected; k_E = 5 voxels). Activations to each condition included regions within superior temporal cortices bilaterally, the left motor cortex, and the anterior cingulate. Visual inspection of these activations would suggests that responses to the PSR and SR conditions are weaker than those to the INIT condition. BOLD responses were submitted to a 1-way repeated measures ANOVA (voxel level P < 0.001, FWE corrected; k_E = 5 voxels), using presentation type as the within-subject factor (INIT, PSR, and SR) and independently of the result shown in Figure 2. When considering only the regressors of the canonical hemodynamic response function, a main effect was observed predominantly within regions of the left hemisphere, including temporal cortices (left and right superior temporal gyrus), frontal cortices (left and right superior frontal gyrus), as well as the cingulate and the medial aspect of the superior frontal gyrus (Fig. 3). Table 1 provides listing of the regions exhibiting a significant main effect as well as the coordinates and β values at the location showing the maximal F value within each cluster. Post hoc analyses revealed that both the
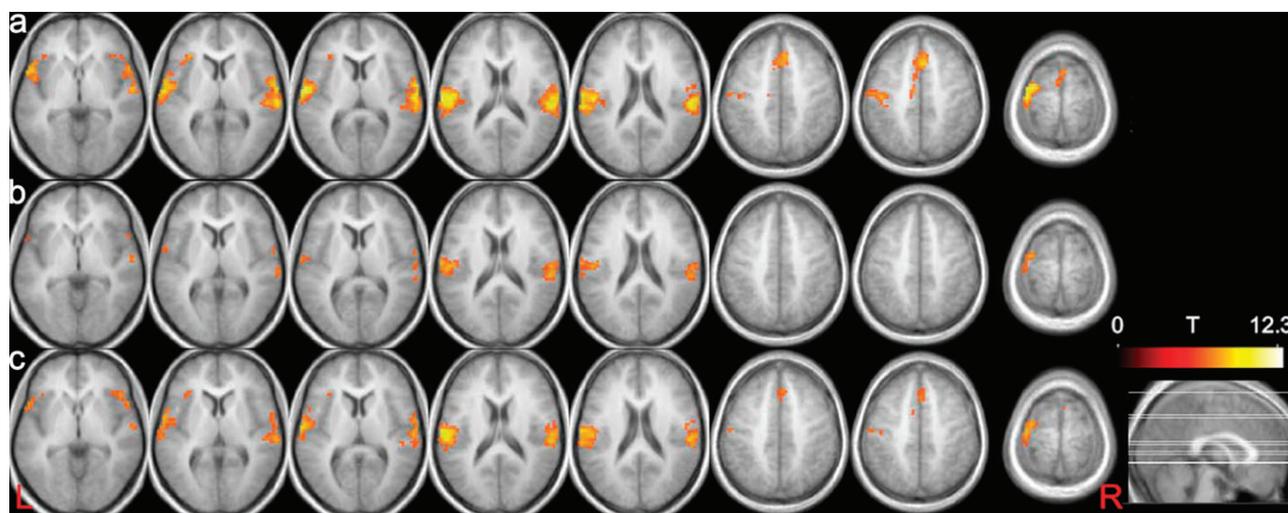

Figure 2. Activation patterns for each stimulus condition. Panels (a–c) depict the contrast of each experimental condition (INIT, PSR, and SR, respectively) versus baseline (P < 0.001, FWE corrected at 5 voxel spatial extent threshold; inset displays the position of the slices) superimposed to the mean T_1 across the 19 subjects.
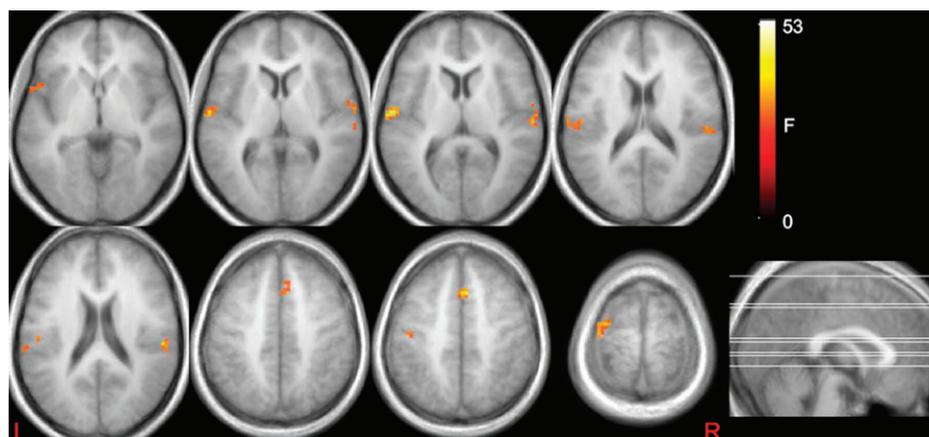

Figure 3. ANOVA results. Axial slices display the main effect ANOVA results (P < 0.001, FWE corrected at 5 voxel spatial extent threshold; inset displays the position of the slices as reported in Table 1 in Montreal Neurological Institute coordinates) superimposed to the mean T_1 across the 19 subjects.

Table 1
Regions showing a main effect in the 1-way ANOVA (voxel-level threshold of $P < 0.001$, FWE corrected; $k_E = 5$ voxels)

Brain area	Brodmann's areas and hemisphere	Number of voxels	Montreal Neurological Institute coordinates			β values (±SEM)			F values
						INIT	PSR	SR	
Superior temporal gyrus	22L, 42L	56	−60	−15	3	13.26 (±1.46)	4.38 (±0.65)	6.60 (±0.76)	45.87
Superior temporal gyrus	21L, 22L	13	−57	3	−6	12.39 (±1.07)	4.37 (±0.46)	6.30 (±0.70)	41.19
Superior temporal gyrus	42L	6	−63	−36	18	12.94 (±1.20)	4.72 (0.88)	6.84 (±1.03)	28.47
Superior frontal gyrus	4L, 6L	67	−30	−12	72	10.15 (±0.79)	4.27 (±0.44)	4.83 (±0.48)	53.28
Postcentral gyrus	3L	5	−45	−21	45	5.96 (±0.84)	1.35 (±0.42)	2.81 (±0.39)	29
Superior and medial frontal gyrus	6R, 32R	50	3	15	48	10.68 (±1.09)	2.11 (±0.62)	4.93 (±0.72)	48.15
Superior temporal gyrus	40R, 42R	14	57	−30	15	11.14 (±1.43)	4.02 (±0.53)	5.66 (±0.78)	36.89
Superior temporal gyrus	21R, 22R, 42R	32	63	−24	6	10.03 (±1.20)	3.88 (±0.57)	5.44 (±0.63)	35.35

Table 2
Regions showing a significant effect of perceptual priming (voxel-level threshold of $P < 0.001$, FWE corrected; $k_E = 5$ voxels)

Brain area	Brodmann's areas and hemisphere	Montreal Neurological Institute coordinates			Number of voxels	T values
Superior temporal gyrus	22L, 42L	−60	−15	3	56	9.11
Superior temporal gyrus	21L, 22L	−57	3	−6	13	7.98
Superior temporal gyrus	42L	−63	−33	15	6	7.79
Precentral and superior frontal gyrus	4L, 6L	−39	−15	66	67	8.79
Postcentral gyrus	3L	−45	−21	45	5	7.02
Cingulate and medial superior frontal gyrus	6R, 8R, 32R	3	12	48	50	9.25
Superior temporal gyrus	40R, 42R	57	−30	15	14	8.19
Superior temporal gyrus	22R, 42R	63	−24	3	32	8.10

Table 3
Regions showing a significant effect of semantic priming (voxel-level threshold of $P < 0.001$, FWE corrected; $k_E = 5$ voxels)

Brain area	Brodmann's areas and hemisphere	Montreal Neurological Institute coordinates			Number of voxels	T values
Superior temporal gyrus	22L, 42L	−60	−15	3	56	8.18
Superior temporal gyrus	21L, 22L	−57	3	−6	13	7.26
Superior temporal gyrus	42L	−63	−33	15	6	6.77
Precentral and superior frontal gyrus	4L, 6L	−30	−12	72	64	9.33
Postcentral gyrus	3L	−42	−24	45	5	5.95
Cingulate and medial superior frontal gyrus	6R, 8R, 32R	3	12	48	46	7.79
Superior temporal gyrus	40R, 42R	57	−30	15	14	7.55
Superior temporal gyrus	22R, 42R	66	−12	0	32	7.28

PSR and SR conditions resulted in repetition suppression ($P < 0.001$, FWE corrected; $k_E = 5$ voxels) within a nearly identical set of brain regions, though the precise coordinates of statistical maxima occasionally varied within a given region (Tables 2 and 3). Direct contrast of the PSR and SR conditions provided no evidence that responses to the different types of repetitions differed (see Table 1), suggesting that both perceptual and semantic priming elicit similar patterns of repetition suppression (for a prototypical β profile, see Fig. 4). No areas showed enhanced activity for stimulus repetitions of either variety relative to initial sound presentation. In addition, there was no evidence for correlations between the extent of repetition suppression and the magnitude of reaction time facilitation (all correlation coefficients < 0.39). Finally, a 1-way ANOVA using presentation type as the within-subject factor and involving only the reaction time regressor did not yield any significant effects (voxel level $P < 0.001$, FWE

corrected; $k_E = 5$ voxels), supporting the conclusion that the present repetition suppression effects were not modulated by reaction time.

A final set of analyses tested for correlations between repetition suppression effects in response to the PSR and SR conditions, while controlling for the influence of the INIT condition (i.e., partial correlations were calculated). Among the 8 clusters that exhibited repetition suppression, significant correlations between effects induced by PSR and SR were observed within 3 regions in the left hemisphere: BA22/42 ($r_{15} = 0.536$; $P = 0.027$), BA42 ($r_{15} = 0.484$; $P = 0.049$), and BA4/6 ($r_{15} = 0.599$; $P = 0.011$).

Discussion

Representations of environmental sounds are subject to plasticity both when the initial and repeated stimuli are identical exemplars as well as when they are physically different but refer to a common referent object (and are thus semantically related). Reaction times were significantly faster for repeated than initial presentations both when an identical exemplar was used (i.e., the PSR condition) and when different exemplars of the same referent object were used (i.e., the SR condition). There was no evidence that the magnitude of the reaction time facilitation differed between perceptual and semantic priming. Nor was there a systematic relationship (correlation) between the magnitude of one and that of the other (Fig. 1). Neurophysiologically, the present fMRI results demonstrate that repetitions of acoustic and/or semantic features produce suppressed BOLD responses within overlapping brain regions ascribed to the auditory "what" pathway (Romanski et al. 1999; Kaas and Hackett 2000) that included not only auditory association cortices but also premotor, prefrontal, and cingulate cortices. Importantly, the present fMRI results also show that the degree of repetition suppression was equivalent both when the acoustic and semantic features repeated and also when only the latter repeated. In this regard, there was no evidence of either a distinct mechanism or network of brain regions mediating semantic priming. Finally, there was no evidence for a systematic relationship between behavioral and neurophysiologic measures of priming, leaving open the issue of a direct causal relationship between neurophysiologic and behavioral manifestations of repetition priming within the auditory modality. Collectively, our results suggest that repetition priming with sounds of environmental objects involves at least minimal access to semantic attributes, as there were similar repetition suppression effects across stimulus conditions.
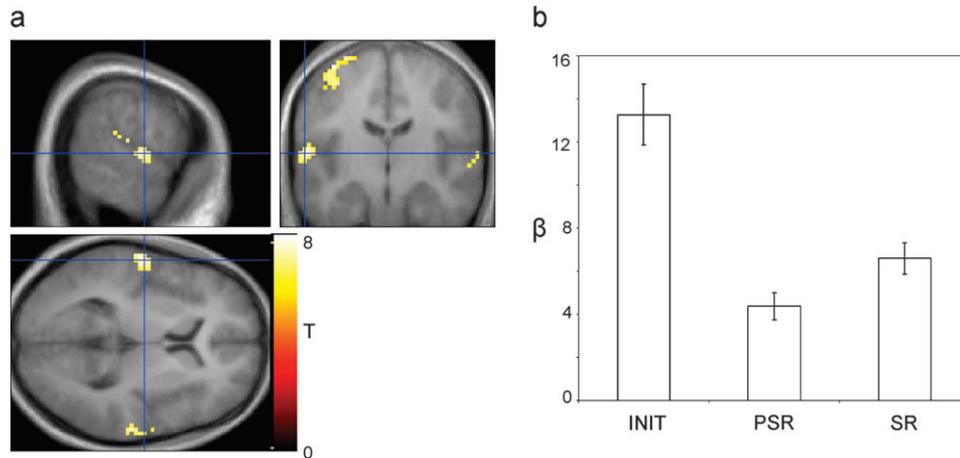
Figure 4. Follow-up contrast results. A cluster in the left superior temporal gyrus (BA22) exhibited reduced activation following repetition of acoustically identical objects (PSR condition; perceptual priming) and for semantically related objects (SR condition; semantic priming); the latter of which did not significantly differ. Panel (*a*) displays the location of these repetition suppression effects. Panel (*b*) displays the mean (SEM indicated) β values at the coordinates indicated by the crosshairs in panel (*a*).

Evidence for Exemplar-Independent Priming

Prior research has provided conflicting results concerning the exemplar independence of behavioral priming with sounds of objects (Stuart and Jones 1995; Chiu 2000). Chiu (2000) suggested that exemplar-independent (i.e., semantic) priming may depend on a lack of (or at least reduced) perceptual discriminability between the exemplars. Our findings indicate otherwise. Exemplar-independent priming was observed despite participants readily discriminating between the exemplars themselves (see Supplementary Material). Plus, our acoustic analyses would indicate that the spectrotemporal difference between different exemplars of the same object is comparable to that between sounds of different objects (Supplementary Fig. S2). However, we should note that while we accounted for shifts in time, we cannot wholly exclude the possibility that there are similar features between different exemplars of the same object that are shuffled in their ordering. In addition, the present study benefited from using a continuous categorization task with initial and repeated presentations intermixed within a single block of trials, thereby conferring the advantages of an event-related design that controls for modulations in attention/arousal and the interval between presentation types. Likewise, the stimuli used in the present study were all of the same, relatively short duration (500 ms) and were controlled across experimental conditions in terms of their spectrogram and HNR, precluding low-level processing differences as an explanation for our effects.

This behavioral evidence for exemplar-independent priming provides some insights on the potential level(s) at which object representations may be reaccessed and/or subject to plasticity. Because prior studies have either failed to demonstrate semantic priming (Chiu 2000) or because perceptual priming effects were localized to superior temporal cortices (Bergerbest et al. 2004), the interpretation was that behavioral and neurophysiologic effects followed from plasticity in analyzing the stimulus features themselves (i.e., their acoustics), rather than in accessing the referent concept (i.e., the object representation in the brain). The present evidence of semantic priming is an indication that plasticity also occurs beyond the level of simple perceptual analysis.

Repetition Suppression and Semantic Representations of Environmental Objects

Repetition suppression within nonprimary superior temporal regions has been previously documented in studies of perceptual priming (Bergerbest et al. 2004; Murray et al. 2008) and adaptation (Ahveninen et al. 2006; Altmann et al. 2007, 2008). Superior temporal, premotor, and prefrontal cortices are believed to be part of an auditory what network involved in sound recognition (Romanski et al. 1999; Kaas and Hackett 2000; Tian et al. 2001; Hall et al. 2002; Lewis et al. 2005; Viceic et al. 2006). While these regions can also be affected by factors such as attention, executive functions, decision making, etc. (Talati and Hirsch 2005), our inclusion of demeaned RT as an additional regressor (see also Desai et al. 2006) allowed us to parcel out the relative contribution of such factors to the present priming effects. That is, the present effects would appear to be specific to stimulus processing rather than to such higher order factors. Moreover, our prior work led to the proposition of a spatiotemporal model of auditory object discrimination. It involves initial categorization within the right BA22 at ~70-ms poststimulus onset. This is followed by access to as well as plasticity in more associative semantic representations within left BA22 starting at ~150-ms poststimulus onset (Murray et al. 2006, 2008). This is in turn followed at ~300 ms by modulations within premotor and prefrontal cortices as a function of associated action representations (De Lucia et al. 2009; Murray and Spierer 2009). The present findings are in solid agreement with this model in that a widely distributed network of brain regions appears to be involved in processing and discriminating between sounds of environmental objects (though some regions may be involved in specific functions at specific latencies and in other functions at subsequent latencies). More generally, they suggest that perceptual priming reflects plasticity at an intermediate level of sound processing possibly relying on the fine analysis of fast spectrotemporal changes (Zatorre 2001; Altmann et al. 2007; Zaehle et al. 2009) and/or rudimentary associative semantic features (Lewis et al. 2005; Murray et al. 2006, 2008). Additional prefrontal and premotor regions were also identified that have been ascribed to both the auditory what pathway (Gifford et al.

2005; Romanski et al. 2005; Cohen et al. 2006; Russ et al. 2008) as well as to an auditory–visual mirror neuron system (Kohler et al. 2002; Keysers et al. 2003; Pizzamiglio et al. 2005; Tettamanti et al. 2005; Gazzola et al. 2006; Lahav et al. 2007; Galati et al. 2008; De Lucia et al. 2009). More germane is that this network of regions exhibits sensitivity to whether or not both the perceptual and semantic features of an object are repeated. Such results support a general model of auditory object recognition that involves a widely distributed network of brain regions (e.g., Lewis et al. 2005; Altmann et al. 2007; Doehrmann et al. 2008). The notion of a widely distributed network mediating the present repetition effects is further supported by the analysis of the correlation in β estimates between the PSR and SR conditions, where significant positive correlations were observed within BA22, BA42, and BA4/6 of the left hemisphere. One possibility is that these regions are playing a common role in both types of repetition suppression.

Divergence with Prior Research

Several contrasts with prior results are also noteworthy. Although repetition suppression was observed in the present study within BA22 and elsewhere (Table 1), our results contrast with those of Bergerbest et al. (2004) in several respects. These authors also obtained repetition enhancement effects predominantly within the fusiform gyrus and precuneus bilaterally, which are areas typically implicated in high-level visual object processes, and also within the left middle frontal gyrus (BA10). Likewise, while Bergerbest et al. (2004) observed significant correlations between behavioral and hemodynamic indices of repetition suppression across a wide network of regions, no such effects were observed either here or in our prior electrical neuroimaging study (Murray et al. 2008). It may thus be the case that shifts in attention/arousal, linked to using a blocked design, mediated the correlations that they observed. By contrast, the event-related design of the present study precludes such an explanation. Moreover, we found no evidence for a significant correlation between behavioral priming and repetition suppression in any identified brain region for either type of priming. Plus, it is important to recall that the inclusion of demeaned RTs as a regressor allowed us to rule out the role of general task effects that were unspecific to priming as contributing to any observed correlations between behavioral priming and repetition suppression. Still, determining a causal link between modulated brain activity and behavioral indices of priming, as has been done for visual stimuli (Wig et al. 2005) will be an important direction for future research. Nonetheless, one common conclusion supported by our work and that of Bergerbest et al. (2004) is that networks mediating priming with sounds of objects are distinct from those involved in similar effects with auditory linguistic stimuli (Buckner et al. 2000; Badgaiyan et al. 2001; Grill-Spector et al. 2006; Raposo et al. 2006). While the former relies principally on regions within the superior temporal cortex, the latter has been reported to rely on extrastriate cortices and to overlap with regions involved in visual (linguistic) priming.

Repetition Suppression as a Form of Plasticity

Repetition suppression may be considered another example of cortical plasticity that may reflect a similar underlying comparison mechanism as the auditory mismatch negativity (MMN). The auditory MMN is a differential brain response to rare or deviant stimuli within a stimulus series (e.g., Näätänen et al. 2005), which is in turn considered to reflect the current stimulus' access to and comparison with a perceptual or memory trace for the consistencies in the stimulus series. The MMN can be elicited on the one hand by changes in low-level acoustic features (e.g., pitch, duration, or location) and on the other hand by alterations in more complex stimulus features, semantic attributes, and arbitrary patterns (e.g., Näätänen et al. 2001). The MMN to such changes in features typically manifests as a signal increase, rather than as repetition suppression (as in the current study), and cannot be explained by adaptation or habituation of sensory components (e.g., Näätänen et al. 2005; for evidence for distinct laminar profiles for sensory processing and MMN, see also Javitt et al. 1996). It is noteworthy that both the MMN and the present repetition effects involve the superior temporal and prefrontal cortices (e.g., Giard et al. 1990; Opitz et al. 2002; Liebenthal et al. 2003). Thus, there is a degree of similarity in terms of recruited neuronal circuitry (at least on a macroscopic level). More recently, it has been proposed that both the MMN and repetition suppression may follow from cortical responses being based on predictive coding of stimuli (Friston 2005; Garrido et al. 2009). Establishing a more direct neurophysiologic link between mechanisms of repetition suppression and MMN generation will require additional experimentation and modeling and will undoubtedly provide insights concerning auditory sensory processing, memory formation/retrieval, and decision making. Nonetheless, the present findings already situate both sets of phenomena within regions of the superior temporal cortex (and elsewhere), further supporting the conceptualization of these regions as a "computational hub" in audition (Griffiths and Warren 2002).

Conclusion

In conclusion, by documenting both perceptual and semantic priming of auditory objects, our results support a model of auditory object processing that is susceptible to plasticity. Moreover, that the loci and magnitude of repetition suppression were insensitive to whether or not both perceptual and semantic information was repeated is suggestive of a degree of acoustically independent semantic analysis in how object representations are maintained and retrieved.

Supplementary Material

Supplementary material can be found at: http://www.cercor. oxfordjournals.org/.

References

Aeschlimann M, Knebel JF, Murray MM, Clarke S. 2008. Emotional pre-eminence of human vocalizations. Brain Topogr. 20:239-248.

Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG, et al. 2006. Task-modulated "what" and "where" pathways in human auditory cortex. Proc Natl Acad Sci USA. 103:14608-14613.

Altmann CF, Doehrmann O, Kaiser J. 2007. Selectivity for animal vocalizations in the human auditory cortex. Cereb Cortex. 17:2601-2608.

Altmann CF, Henning M, Döring MK, Kaiser J. 2008. Effects of feature-selective attention on auditory pattern and location processing. Neuroimage. 41:69-79.

Badgaiyan RD, Schacter DL, Alpert NM. 2001. Priming within and across modalities: exploring the nature of rCBF increases and decreases. Neuroimage. 13:272-282.

Bergerbest D, Ghahremani DG, Gabrieli JDE. 2004. Neural correlates of auditory repetition priming: reduced fMRI activation in the auditory cortex. J Cogn Neurosci. 16:966-977.

Bregman AS. 1990. Auditory scene analysis. Cambridge (MA): MIT Press.

Buckner RL, Koutstaal W, Schacter DL, Rosen BR. 2000. Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. Brain. 123:620-640.

Chiu CYP. 2000. Specificity of auditory implicit and explicit memory: is perceptual priming for environmental sounds exemplar specific? Mem Cognit. 28:1126-1139.

Cohen YE, Hauser MD, Russ BE. 2006. Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. Biol Lett. 2:261-265.

De Lucia M, Camen C, Clarke S, Murray MM. 2009. The role of actions in auditory object discrimination. Neuroimage. 48:476-486.

Desai R, Conant LL, Waldron E, Binder JR. 2006. FMRI of past tense processing: the effects of phonological complexity and task difficulty. J Cogn Neurosci. 18:278-297.

Doehrmann O, Naumer MJ, Volz S, Kaiser J, Altmann CF. 2008. Probing category selectivity for environmental sounds in the human auditory brain. Neuropsychologia. 46(11):2776-2786.

Friston K. 2005. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 360:815-836.

Friston KJ, Jezzard P, Turner R. 1994. Analysis of functional MRI time-series. Hum Brain Mapp. 1:153-171.

Galati G, Committeri G, Spitoni G, Aprile T, Di Russo F, Pitzalis S, Pizzamiglio L. 2008. A selective representation of the meaning of actions in the auditory mirror system. Neuroimage. 40:1274-1286.

Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Baldeweg T, Friston KJ. 2009. Repetition suppression and plasticity in the human brain. Neuroimage. 48:269-279.

Gazzola V, Aziz-Zadeh L, Keysers C. 2006. Empathy and the somatotopic auditory mirror system in humans. Curr Biol. 16:1824-1829.

Giard MH, Perrin F, Pernier J, Bouchet P. 1990. Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. Psychophysiology. 27:627-640.

Gifford GW, 3rd, MacLean KA, Hauser MD, Cohen YE. 2005. The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. J Cogn Neurosci. 17:1471-1482.

Griffiths TD, Warren JD. 2002. The planum temporale as a computational hub. Trends Neurosci. 25:348-353.

Grill-Spector K, Henson RN, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci. 10:14-23.

Hall DA, Johnsrude IS, Haggard MP, Palmer AR, Akeroyd MA, Summerfield AQ. 2002. Spectral and temporal processing in human auditory cortex. Cereb Cortex. 12:140-149.

Henson RNA. 2003. Neuroimaging studies of priming. Prog Neurobiol. 70:53-81.

Holmes AP, Friston KJ. 1998. Generalisability, random effects and population inference. Neuroimage. 7:S754.

Javitt DC, Steinschneider M, Schroeder CE, Arezzo JC. 1996. Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. Proc Natl Acad Sci USA. 93:11962-11967.

Kaas JH, Hackett TA. 2000. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci USA. 97:11793-11799.

Keysers C, Kohler E, Umilta MA, Nanetti L, Fogassi L, Gallese V. 2003. Audiovisual mirror neurons and action recognition. Exp Brain Res. 153:628-636.

Knebel JF, Toepel U, Huldry J, le Coutre J, Murray MM. 2008. Generating controlled image sets in cognitive neuroscience research. Brain Topogr. 20:284-289.

Kohler E, Keysers C, Umilta MA, Fogassi L, Gallese V, Rizzolatti G. 2002. Hearing sounds, understanding actions: action representation in mirror neurons. Science. 297:846-848.

Lahav A, Saltzman E, Schlaug G. 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. J Neurosci. 27:308-314.

Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA. 2005. Distinct cortical pathways for processing tool versus animal sounds. J Neurosci. 25:5148-5158.

Liebenthal E, Ellingson ML, Spanaki MV, Prieto TE, Ropella KM, Binder JR. 2003. Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. Neuroimage. 19(4):1395-1404.

Manly BF. 1991. Randomization and Monte Carlo methods in biology. London (UK): Chapman & Hall.

Murray MM, Camen C, Andino SLG, Bovet P, Clarke S. 2006. Rapid brain discrimination of sounds of objects. J Neurosci. 26:1293-1302.

Murray MM, Camen C, Spierer L, Clarke S. 2008. Plasticity in representations of environmental sounds revealed by electrical neuroimaging. Neuroimage. 39:847-856.

Murray MM, Spierer L. 2009. Auditory spatio-temporal brain dynamics and their consequences for multisensory interactions in humans. Hear Res. doi: 10.1016/j.heares.2009.04.022.

Näätänen R, Jacobsen T, Winkler I. 2005. Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. Psychophysiology. 42:25-32.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. 2001. "Primitive intelligence" in the auditory cortex. Trends Neurosci. 24:283-288.

Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schröger E. 2002. Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. Neuroimage. 15:167-174.

Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S, Di Russo F. 2005. Separate neural systems for processing action- or non-action-related sounds. Neuroimage. 24:852-861.

Raposo A, Moss HE, Stamatakis EA, Tyler LK. 2006. Repetition suppression and semantic enhancement: an investigation of the neural correlates of priming. Neuropsychologia. 44:2284-2295.

Romanski LM, Averbeck BB, Diltz M. 2005. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. J Neurophysiol. 93:734-747.

Romanski LM, Bates JF, Goldman-Rakic PS. 1999. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol. 403:141-157.

Russ BE, Ackelson AL, Baker AE, Cohen YE. 2008. Coding of auditory-stimulus identity in the auditory non-spatial processing stream. J Neurophysiol. 99:87-95.

Schacter DL, Buckner RL. 1998. Priming and the brain. Neuron. 20:185-195.

Schacter DL, Dobbins IG, Schnyer DM. 2004. Specificity of priming: a cognitive neuroscience perspective. Nat Rev Neurosci. 5:853-862.

Stuart GP, Jones DM. 1995. Priming the identification of environmental sounds. Q J Exp Psychol A. 48:741-761.

Talati A, Hirsch J. 2005. Functional specialization within the medial frontal gyrus for perceptual go/no-go decisions based on "what," "when," and "where" related information: an fMRI study. J Cogn Neurosci. 17:981-993.

Tettamanti M, Buccino G, Saccuman MC, Gallese V, Danna M, Scifo P, Fazio F, Rizzolatti G, Cappa SF, Perani D. 2005. Listening to action-related sentences activates fronto-parietal motor circuits. J Cogn Neurosci. 17:273-281.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. 2001. Functional specialization in rhesus monkey auditory cortex. Science. 292:290-293.

Viceic D, Fornari E, Thiran JP, Maeder PP, Meuli R, Adriani M, Clarke S. 2006. Human auditory belt areas specialized in sound recognition: a functional magnetic resonance imaging study. Neuroreport. 17:1659-1662.

Wig GS, Grafton ST, Demos KE, Kelley WM. 2005. Reductions in neural activity underlie behavioral components of repetition priming. Nat Neurosci. 8:1228-1233.

Zaehle T, Jancke L, Herrmann CS, Meyer M. 2009. Pre-attentive spectro-temporal feature processing in the human auditory system. Brain Topogr. 22:97-108.

Zatorre RJ. 2001. Neural specializations for tonal processing. Ann N Y Acad Sci. 930:193-210.