Behavioral/Systems/Cognitive

Spoken Word Memory Traces within the Human Auditory Cortex Revealed by Repetition Priming and Functional Magnetic Resonance Imaging

Pierre Gagnepain, Gael Chételat, Brigitte Landeau, Jacques Dayan, Francis Eustache, and Karine Lebreton
Inserm-Ecole Pratique des Hautes Etudes-Université de Caen/Basse-Normandie, Unité 923, Groupement d'Intérêt Public Cyceron, Centre Hospitalier Universitaire Côte de Nacre, F-14074 Caen, France

Previous neuroimaging studies in the visual domain have shown that neurons along the perceptual processing pathway retain the physical properties of written words, faces, and objects. The aim of this study was to reveal the existence of similar neuronal properties within the human auditory cortex. Brain activity was measured using functional magnetic resonance imaging during a repetition priming paradigm, with words and pseudowords heard in an acoustically degraded format. Both the amplitude and peak latency of the hemodynamic response (HR) were assessed to determine the nature of the neuronal signature of spoken word priming. A statistically significant stimulus type by repetition interaction was found in various bilateral auditory cortical areas, demonstrating either HR suppression and enhancement for repeated spoken words and pseudowords, respectively, or word-specific repetition suppression without any significant effects for pseudowords. Repetition latency shift only occurred with word-specific repetition suppression in the right middle/posterior superior temporal sulcus. In this region, both repetition suppression and latency shift were related to behavioral priming. Our findings highlight for the first time the existence of long-term spoken word memory traces within the human auditory cortex. The timescale of auditory information integration and the neuronal mechanisms underlying priming both appear to differ according to the level of representations coded by neurons. Repetition may "sharpen" word-nonspecific representations coding short temporal variations, whereas a complex interaction between the activation strength and temporal integration of neuronal activity may occur in neuronal populations coding word-specific representations within longer temporal windows.

Key words: repetition priming; spoken word; auditory cortex; fMRI; repetition suppression; repetition enhancement

Introduction

Even in considerably degraded perceptual situations, the human brain can efficiently recognize words, faces, sounds and objects. To achieve a coherent percept, the brain is thought to draw on cortical memory systems located in perceptual processing areas that encode the physical properties of stimuli but not their meaning (Tulving and Schacter, 1990).

Neuroimaging studies using the long-term repetition priming paradigm have supported this idea with regard to the visual system, showing that some neuronal populations along the visual processing pathway are sensitive to the repetition of both familiar and unfamiliar materials (Henson, 2000; Thiel et al., 2002; Fiebach et al., 2005). Repetition of familiar material usually entails a decrease in neuronal activity, referred to as repetition suppression, which may rely on a sharpened representation composed of neurons that code only key stimulus features (Desimone, 1996;

Wiggs and Martin, 1998; Grill-Spector et al., 2006). Repetition suppression has been found in various visual areas (for review, see Schacter et al., 2004) coding written words (Schott et al., 2006), faces (Eger et al., 2005), objects (Vuilleumier et al., 2002) or line drawings (Lebreton et al., 2001). Repetition of unfamiliar material, however, sometimes leads to an increase in cortical activity (Schacter et al., 1995a; Habeck et al., 2006), referred to as repetition enhancement, which may reflect the formation of new cell assemblies (Henson, 2000; Fiebach et al., 2005).

Concerning auditory repetition priming, neuroimaging studies have previously failed to confirm the involvement of the auditory cortex (Badgaiyan et al., 1999, 2001; Carlesimo et al., 2004; Orfanidou et al., 2006) (for neuropsychological evidence, see also Swick et al., 2004). However, a large body of behavioral evidence suggests that this phenomenon is indeed mediated by perceptual representations (Schacter and Church, 1992; Church and Schacter, 1994), i.e., by the auditory cortex. The tasks used in previous neuroimaging studies may have minimized the extent to which perceptual information is processed. Moreover, these studies assessed repetition-related changes in terms of the mean neuronal firing rate, rather than temporal neuronal firing. This dimension, which is likely to be related to the latency of the hemodynamic response (HR) (Henson et al., 2002), measured by functional magnetic resonance imaging (fMRI) may, however,

play an important role in repetition priming, as proposed by the facilitation/accumulation models (Henson and Rugg, 2003; Grill-Spector et al., 2006; James et al., 2006).

Accordingly, this study sought to assess the existence of neuronal populations coding the physical properties of spoken words in the auditory cortex and to enhance our understanding of the neuronal mechanisms underlying auditory repetition priming. Brain activity was measured using fMRI while subjects performed a lexical decision task on primed and unprimed words and pseudowords, heard in an acoustically degraded format to optimize the recruitment of these hypothetical perceptual representations (Luce and Lyons, 1998). Moreover, we analyzed both the amplitude and latency of the HR to determine the nature of the neuronal signature of spoken word priming.

Materials and Methods

Participants. Eighteen right-handed French men aged 20–29 years old (mean, 25.1; SD, 2.2) were paid to take part in the study. They had no reported history of neurological, medical, speech or hearing disorders, had had at least 14 years of schooling and presented normal MRI structural images. The project was approved by the regional ethics committee and all subjects gave their written consent.

Stimuli. The material consisted of six lists, each containing 40 words or 40 pseudowords. Word lists were paired according to frequency, concreteness, gender, phonological point of unity, consonantal structure (e.g., CVCCV), and number of phonemes, syllables, and phonological neighbors. The pseudoword lists were created and matched to the word lists according to the number of syllables and phonemes, and the consonantal structure. Particular care was taken to ensure that none of these pseudowords sounded like real words. Two additional lists of words and pseudowords were also created for the training phases (see below). All of the items, spoken by a female voice, were deliberately degraded for the functional session using Adobe (San Jose, CA) Audition audio processing software by applying a Butterworth third-order low-pass filter with a frequency cutoff point of 1000 Hz (Fig. 1b). The filter let through all of the frequencies between 0 and 1000 Hz in their original intensity. However, the frequencies between ~1000 and 5500 Hz were reduced by between 0 and 45 dB in an exponential curve (0 dB for 1000 Hz and the frequencies just above and 45 dB for the frequencies bordering on 5500 Hz). Twenty consonantal sounds used during the phonemic study phase (see below) were also recorded.

Procedure. The repetition priming task (Fig. 1a,b) was composed of two distinct phases, each preceded by a training phase and separated by a 45 min interval during which acquisition of structural images was performed. During the first phase (study phase), occurring outside the MRI scanner, subjects performed the phonemic processing of 120 word primes and 120 pseudowords primes. The subjects were required to indicate via a keyboard whether or not the item contained the consonantal sound that was presented to them after a 500 ms interval. This task sought to induce and encourage the processing of the perceptual information of the speech items through consonant detection. Twenty consonantal sounds were chosen: /b/, /d/, /f/, /g/, /gz/, /j/, /k/, /ks/, /l/, /m/, /n/, /p/, /r/, /s/, /t/, /v/, /w/, /z/, /ʃ/, and/ɲ/. Positive (consonantal sound present in
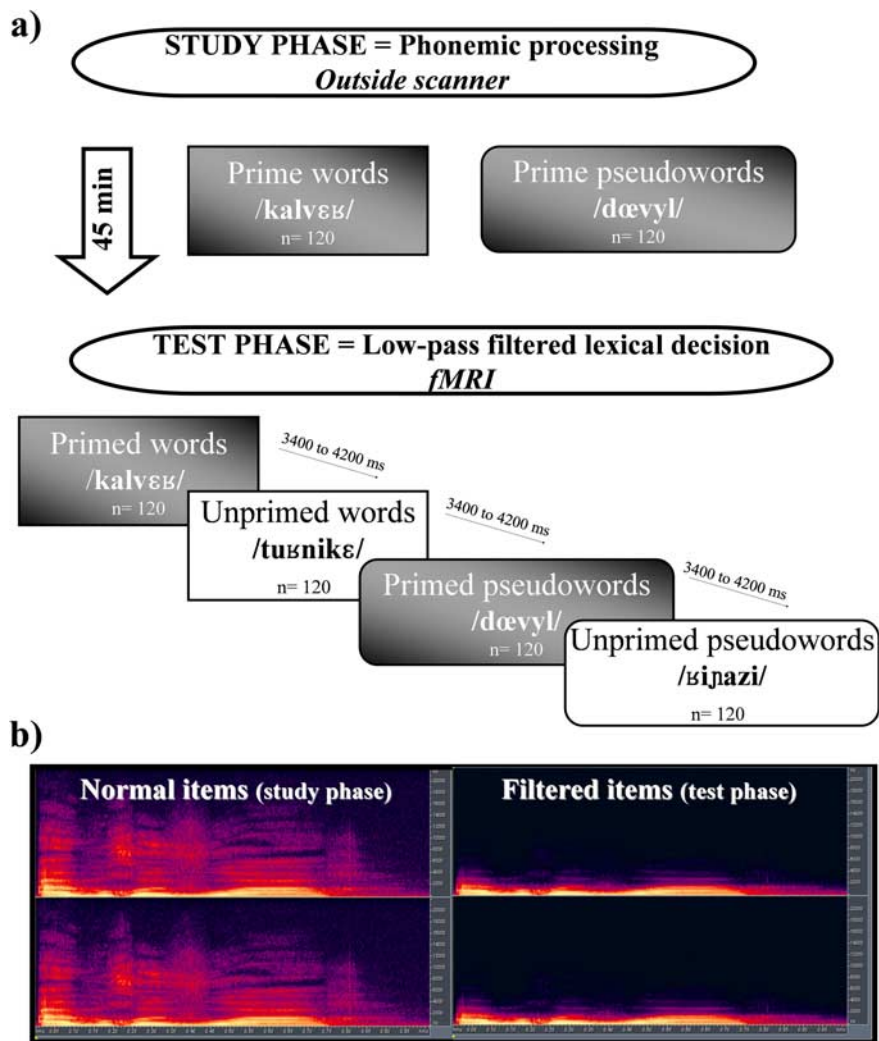

Figure 1. *a,* Repetition priming paradigm comprising (1) a study phase where 120 spoken word and pseudoword primes were presented outside the scanner during a phonemic processing task and (2) a test phase consisting of a low-pass-filtered lexical decision task on primed and unprimed spoken words and pseudowords during brain activity recording. The graphemic translation of the stimulus examples provided in the figure is as follows: calvaire, tourniquet, deuvul, rignazi. *b,* Spectral view of normal (study phase) and low-pass-filtered (test phase) items (*x* = time, *y* = frequency; colors indicate energy).

the item) and negative responses (not present) were counterbalanced across items and subjects and, as far as possible, positive responses were placed equally often at the beginning, middle and end of the items. A distracter task, which consisted in counting backwards in threes for 2 min and was intended to prevent maintenance or elaboration of the primed items, was performed by subjects immediately after this first phase. The real objectives of the study were not disclosed to the subjects, who were simply told that they were taking part in a study of auditory perception.

The second phase (test phase) was performed during fMRI acquisition and was divided into three functional sessions. Each session featured 40 primed and unprimed words and pseudowords, leading to a total of 160 stimuli per session. Subjects performed a low-pass-filtered lexical decision task, in which they had to indicate as fast as possible using their right-hand whether or not the items they had heard corresponded to a word belonging to the French language. Subjects were also instructed to close their eyes to better focus on the task. For each trial, the items started without any cue and subjects' reaction times (RTs) were measured from the onset of the stimuli. Note that all item lists matched in term of mean stimulus duration avoiding any confusion with RT priming effects. Among the various possible combinations of the six lists of words and pseudowords, 18 were selected in a pseudorandom way, so that (1) the primed and unprimed status of the lists and (2) the functional session in which the lists were used were counterbalanced across the subjects. Fur-

thermore, this counterbalancing was performed so that comparisons between primed and unprimed conditions for individual subjects never involved the same item lists. All trial conditions (word_primed, word_unprimed, pseudoword_primed, pseudoword_unprimed) were presented according to an efficient stochastic design (Friston et al., 1999) and the optimal order for detecting differences between primed and unprimed items was computed using a genetic algorithm (Wager and Nichols, 2003). The interstimulus interval varied between 3400 and 4200 ms (mean, 3800 ms; 1.9 TRs) (see below), ensuring that the HR was sampled approximately every 200 ms over the trials. Items were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA) implemented within the IFIS System Manager (Invivo, Orlando, FL), controlling stimulus delivery by synchronizing each trial with a series of transistor–transistor logic pulses produced during imaging acquisition. Items were delivered via an electrodynamic audio system (MR Confon, Magdeburg, Germany) ensuring an attenuation of scanner noise up to 45 dB.

Image acquisition. All images were acquired using the Philips (Eindhoven, The Netherlands) 3T system. Blood oxygen level-dependent (BOLD) images were collected using a T2*-weighted fast field echo echoplanar imaging sequence [$64 \times 64 \times 31$; $3.5 \times 3.5 \times 3.5$ mm^3; field of view (FOV), 224 mm; echo time (TE), 35 ms; flip angle, 80°, repetition time (TR), 2000 ms]. The slices were acquired in an interleaved ascending direction. The 936 functional volumes were collected during three functional sessions of 312 volumes each, where the first six volumes were discarded to allow for equilibration effects. T1-weighted structural images were also acquired ($256 \times 256 \times 180$; $1 \times 1 \times 1$ mm^3; TE, 4.6 ms; flip angle, 20; FOV, 256; TR, 20 ms).

fMRI data processing. Data were analyzed using statistical parametric mapping software (SPM5; Wellcome Department of Cognitive Neurology, Institute of Neurology, London, England). During preprocessing, images were first corrected for slice acquisition temporal delay, before being spatially realigned to correct for movement. Images were then normalized using parameter derived from the normalization of individual gray-matter T1 images to the T1 template of the Montreal Neurological Institute (MNI), and spatially smoothed using an 8 mm full-width at half-maximum Gaussian kernel.

In the initial analysis, the resulting preprocessed time series were high-pass filtered in each voxel to 1/128 Hz and globally normalized with scaling option. A general linear model was applied to each participant by creating a δ function with stimulus onset for each condition of interest (word_primed, word_unprimed, pseudoword_primed, pseudoword_unprimed) plus an additional condition for incorrect responses that was then convolved with the canonical hemodynamic response function (HRF) and its time and dispersion derivatives (Friston et al., 1998). Covariates of no interest included the six realignment parameters to account for motion artifacts. Regressor parameters were then estimated including a correction for the temporal autocorrelation across volumes according to an AR(1) process (Friston et al., 2002). Linear combinations of the canonical HRF (activation maps) and time-derivative estimated parameters were then computed for each condition of interest. Individual activation maps of the main network involved in the task were also computed against the implicit baseline (i.e., constant term of the model) by combining together the HRF canonical parameters corresponding to each condition of interest. Latency maps corresponded to the derivative/canonical ratio for each voxel and condition of interest, which was transformed using a sigmoidal function (for details of the method, see Henson et al., 2002). These latency maps expressed the difference in seconds between the HR and canonical HRF time to peak.

In the second-level analysis, with subjects treated as a random effect, we primarily entered individual main effects of the task into a one-sample *t* test thresholded at a $p < 0.001$ false discovery rate (FDR) (Genovese et al., 2002), corrected for multiple comparisons, to identify the main network involved in the task. Note that this threshold was preferred because it was a reasonable compromise between the presence of many type I errors inherent to a more permissive FDR correction for such large network and an overly conservative threshold provided by the familywise error rate correction. All subsequent analyses were conducted within this network. This procedure allowed us to exclude regions showing a negative "deactivation" BOLD response during group analysis and

to focus latency analyses on regions providing a reasonable match with the canonical HRF (Henson et al., 2002).

Individual activation and latency maps for each condition of interest were entered into two separate factorial ANOVAs implemented in SPM, which used pooled error and correction for nonsphericity to create *t* statistics. These *t* maps were then thresholded at $p < 0.05$ after FDR correction for multiple comparisons (cluster extent thresholded at $k = 5$ voxels). When no voxels survived this correction, a $p < 0.001$ uncorrected threshold was also applied.

For the ANOVA on activation maps implemented in SPM, main effects of item type (words − pseudowords; pseudowords − words) and repetition (primed − unprimed; unprimed − primed) were computed, as well as stimulus type by repetition interaction effects [(words_primed − words_unprimed) − (pseudowords_primed − pseudowords_unprimed); (pseudowords_primed − pseudowords_unprimed) − (words_primed − words_unprimed)]. Within the regions of interest (ROIs) showing significant interaction, the amplitude of the HR was averaged across voxels defined by each SPM cluster for each subject and condition on the basis of the estimated parameters of the canonical HRF, and entered into an ROI-based ANOVA to assess the effect of repetition for each item type. Furthermore, to assess whether repetition in interacting regions also affects latency of the HR, the mean HR latency was also extracted in the same regions and entered into an ROI-based ANOVA. Last, these mean HR amplitude and latency differences computed by subtracting unprimed items from primed ones were entered into four independent ROI-based multiple-regression analyses as predictors of the behavioral priming score variance corresponding to the word priming RT, pseudoword priming RT, word priming accuracy score, or pseudoword priming accuracy score. These behavioral priming scores were computed by subtracting RTs or accuracy scores (proportion of correct responses) of unprimed items from those of primed items.

For the ANOVA performed on the latency maps in SPM, only the main effect of repetition and the interaction effects were computed. As for the activation analysis, both HR amplitude and latency were averaged across voxels corresponding to each SPM cluster that showed significant interaction and entered into ROI-based ANOVAs to assess the impact of repetition. Figure 2 illustrates the organization of these voxel-based and ROI-based second-level analyses.

Results

Behavioral results

Behavioral results are summarized in Figure 3. ANOVAs on RTs and accuracy scores were conducted using stimulus type (word vs pseudoword) and repetition (primed vs unprimed) as repeated measures. Results showed a significant main effect of stimulus type for both RTs and accuracy (acc) scores (RT, $F_{(1,17)} = 34.8$, $p = 0.00002$; acc, $F_{(1,17)} = 49.2$, $p = 0.00002$), a main effect of repetition only for RTs (RT, $F_{(1,17)} = 6.02$, $p = 0.025$; acc, $F_{(1,17)} = 3.2$), and a significant interaction effect for both RTs and accuracy scores (RT, $F_{(1,17)} = 8.5$, $p = 0.01$; acc, $F_{(1,17)} = 10.55$, $p = 0.004$). For both RTs and accuracy scores, a significant repetition effect (RT, $F_{(1,17)} = 11.58$, $p = 0.003$; acc, $F_{(1,17)} = 11.96$, $p = 0.003$) was found for words_primed relative to words_unprimed, but not for pseudowords (RT, $F_{(1,17)} = 0.44$; acc, $F_{(1,17)} = 0.49$).

fMRI results

Main network involved in the task

Regardless of type or repetition, the presentation of stimuli activated a large network ($p < 0.001$ FDR corrected for multiple comparisons), which extended mainly bilaterally, to the middle and superior temporal regions, dorsal and posterior frontal lobes, and cerebellum, but also to the left inferior parietal regions, putamen, pallidum, and thalamus (supplemental Table S1, available at www.jneurosci.org as supplemental material). All subsequent analyses were restricted to this network (see Materials and Methods) reported in Figure 4.

ANOVA on activation maps
Main effects of item type and repetition. Increased activity for words relative to pseudowords was found in the left inferior frontal gyrus (extending to the insula and superior temporal pole), the left posterior middle temporal gyrus (MTG), the left inferior precentral gyrus (extending to the inferior frontal gyrus), and the left pallidum and putamen (supplemental Table S2, available at www.jneurosci.org as supplemental material). Increased activity for pseudowords relative to words was found in the postcentral gyrus (extending to the precentral and inferior parietal gyri), the right cerebellum, the right middle cingulate gyrus (extending to the supplementary motor area), the bilateral superior temporal gyrus (STG), and the left rolandic operculum (extending to the postcentral gyrus) (supplemental Table S2, available at www.jneurosci.org as supplemental material). There were no main effects of repetition in any direction, either with the $p < 0.05$ FDR-corrected or the $p < 0.001$ uncorrected thresholds.

Stimulus type by repetition interactions. No brain regions showed significant effects for the [(words$_{primed}$ − words$_{unprimed}$) − (pseudowords$_{primed}$ − pseudowords$_{unprimed}$)] contrast, either with the corrected or the uncorrected threshold. For the reverse contrast [(words$_{unprimed}$ − words$_{primed}$) − (pseudowords$_{unprimed}$ − pseudowords$_{primed}$)], however, a significant effect was found in the left planum temporale [PT; extending to the posterior superior temporal sulcus (STS) and posterior MTG], the left precentral gyrus (extending to the inferior frontal gyrus), the right Heschl's gyrus (HG; extending to the STG), the left cerebellum, the right anterior STG (extending to temporal superior pole), the right middle/posterior STS, the left temporal superior pole, and the right middle STS. Note that the peak coordinates of the left PT fell within the probability maps described by Westbury et al. (1999). Detailed anatomical and statistical information is set out in Table 1 and functional displays are provided in Figure 5, *a* and *b*.

ROI-based ANOVAs on the mean amplitude of activity were then performed within this set of areas. Note that the cluster found in the left temporal regions ($x = -60$, $y = -36$, $z = 10$) (Table 1) was divided into two distinct ROI covering 90% of the main cluster: the left PT extending to the posterior STS and the left posterior MTG. Statistically significant word repetition suppression and pseudoword repetition enhancement were found in the left PT extending to the posterior STS (words, $F_{(1,17)} = 12.44$, $p = 0.002$; pseudowords, $F_{(1,17)} = 6.98$, $p = 0.017$), the right HG extending to the STG (words, $F_{(1,17)} = 7.43$, $p = 0.014$; pseudowords, $F_{(1,17)} = 5.54$, $p = 0.03$), the left cerebellum (words, $F_{(1,17)} = 3.62$, $p = 0.074$; pseudowords, $F_{(1,17)} = 5.69$, $p = 0.029$), and the left precentral gyrus extending to the inferior frontal gyrus (words, $F_{(1,17)} = 22.83$, $p = 0.0002$; pseudowords, $F_{(1,17)} = 6.52$, $p = 0.02$). Statistically significant word-specific
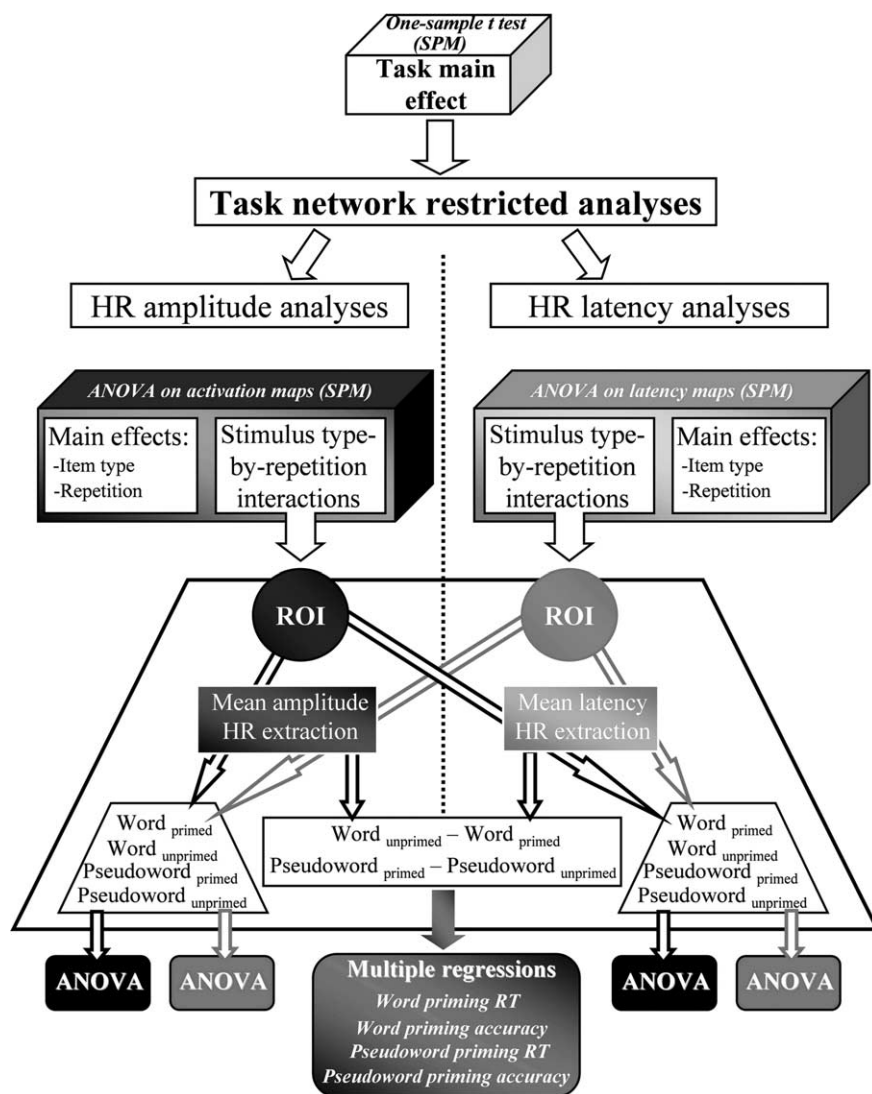

Figure 2. General organization of the voxel-based and ROI-based second-level analyses (see fMRI data processing section for details).

repetition suppression without any repetition effect on pseudowords was found in the right middle/posterior STS ($F_{(1,17)} = 13.6$; $p = 0.006$), right anterior STG ($F_{(1,17)} = 16.87$; $p = 0.0007$), right middle STS ($F_{(1,17)} = 9.13$; $p = 0.008$), left posterior MTG ($F_{(1,17)} = 10.46$; $p = 0.005$), and left temporal superior pole ($F_{(1,17)} = 6.42$; $p = 0.021$). For all of the ROI-based analyses described here, p values were not corrected for multiple comparisons because of the subtlety of changes associated with repetition priming, which may be overlooked using thresholds that are too conservative.

ROI-based ANOVAs on mean HR latency revealed an earlier HR peak for words$_{primed}$ relative to words$_{unprimed}$ in the right middle/posterior STS ($F_{(1,17)} = 5.45$; $p = 0.032$) and left posterior MTG ($F_{(1,17)} = 6.82$; $p = 0.018$). A delayed HR peak was found for pseudowords$_{primed}$ relative to pseudowords$_{unprimed}$ in the right anterior STG ($F_{(1,17)} = 4.9$; $p = 0.04$).

ROI-based multiple-regression analyses showed statistically significant relationships between word priming RT and (1) word-specific repetition suppression in the right middle/posterior STS ($\beta = 0.51$; $R^2 = 35.6\%$; $F_{(1,15)} = 8.28$; $p = 0.01$) (Fig. 5a) and (2) the repetition latency shift in the right middle/posterior STS ($\beta = $
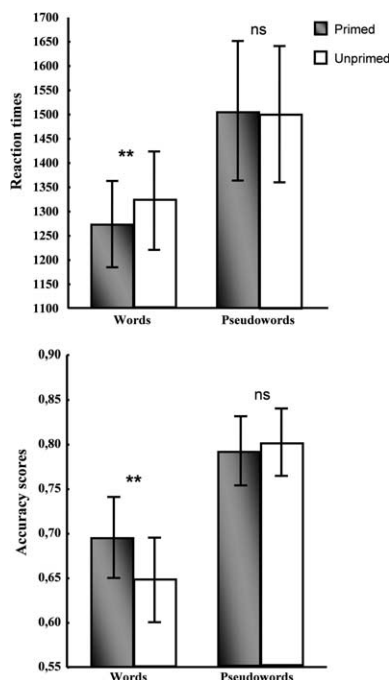
Figure 3. Mean reaction times and accuracy scores (95% confidence intervals represented by error bars) in the low-pass-filtered lexical decision task according to item type (words vs pseudowords) and priming conditions (primed vs unprimed). Significant differences between primed and unprimed items in terms of reaction times and accuracy scores are **$p < 0.005$. ns, Nonsignificant difference.
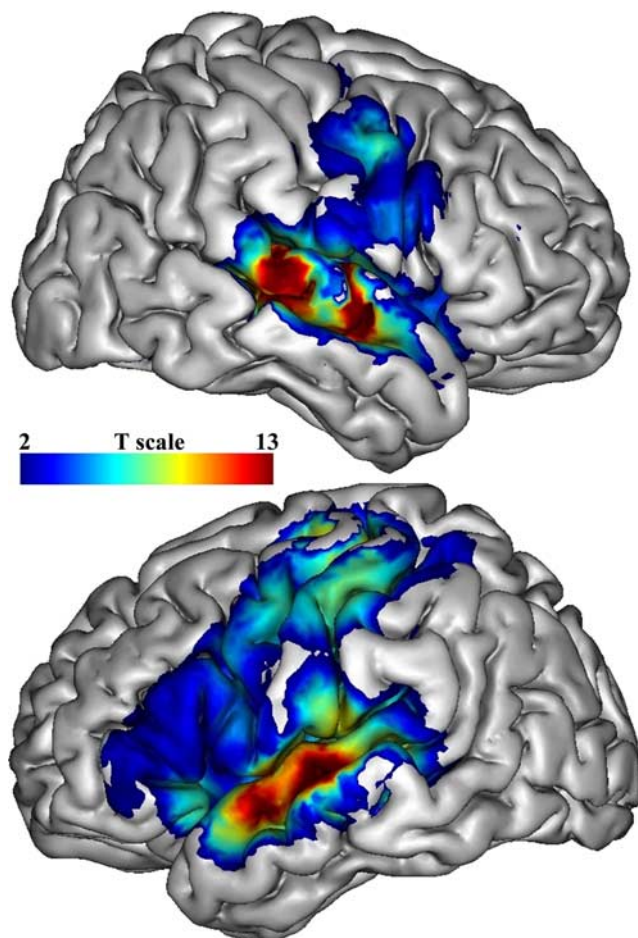

Figure 4. fMRI activations ($p < 0.001$, FDR corrected for multiple comparisons) for the main effect of the task, rendered onto an inflated reconstruction of an MNI template brain. All of the analyses were conducted within this network.

-0.42; $R^2 = 27.3\%$; $F_{(1,15)} = 6.64$; $p = 0.03$) (Fig. 5a) and the left precentral gyrus, extending to the inferior frontal gyrus ($\beta = -0.55$; $R^2 = 28.8\%$; $F_{(1,15)} = 6.05$; $p = 0.026$). No significant relationships were found in any other regression analysis (i.e., with pseudoword priming RT, word priming accuracy score or pseudoword priming accuracy score as dependent variables). Figure 5, a and b, illustrates the main findings of these ROI-based ANOVAs and multiple-regression analyses within the temporal regions.

ANOVA on latency maps
Main effects of repetition. No regions showed a significant earlier HR peak for primed relative to unprimed items, but a delayed HR peak of primed relative to unprimed items was observed in the left anterior STG, extending to the left rolandic operculum at $p < 0.001$ uncorrected ($x = -50$, $y = -6$, $z = 0$; cluster extent = 16; $T_{max} = 3.65$; $p_{uncorrected} = 0.0003$), although it did not survive the FDR-corrected threshold.

Stimulus type by repetition interactions. No brain regions showed significant effects for the [(words$_{primed}$ − words$_{unprimed}$) − (pseudowords$_{primed}$ − pseudowords$_{unprimed}$)] contrast, using either the corrected or the uncorrected threshold. For the reverse contrast [(words $_{unprimed}$ − words $_{primed}$) − (pseudowords $_{unprimed}$ − pseudowords $_{primed}$)], however, a significant effect was found in the right middle/posterior STG (Fig. 5c) at $p < 0.001$ uncorrected ($x = 62$, $y = -32$, $z = 10$; cluster extent = 29; $T_{max} = 4.11$; $p_{uncorrected} = 0.00007$), but no voxels survived the FDR correction.

ROI-based ANOVA on mean HR latency revealed an earlier HR peak in this region for words$_{primed}$ relative to words$_{unprimed}$ ($F_{(1,17)} = 11.22$; $p = 0.0038$) and a delayed HR peak for pseudowords$_{primed}$ relative to pseudowords $_{unprimed}$ ($F_{(1,17)} = 7.93$; $p = 0.012$). No significant differences were found with re-

gard to the mean amplitude of activity for either item type. Note that we were able to confirm that this region did not overlap with interacting regions in the activation analysis by excluding these regions in a complementary analysis. A functional display of the stimulus type by repetition interaction found in the latency analysis is provided in Figure 5c, together with the ROI-based ANOVA results.

Discussion

We used fMRI and repetition priming to explore how the human auditory cortex may code the physical properties of spoken words. We found stimulus type by repetition interactions in various regions of the auditory cortex, characterized either by a suppressed or enhanced HR for repeated spoken words and pseudowords respectively or by word-specific repetition suppression (Fig. 5a,b). Word-specific repetition suppression was associated with a repetition latency shift in the right middle/posterior STS, and these two neuronal priming indexes were related to behavioral priming (Fig. 5a). Last, a stimulus type by repetition interaction in terms of HR latency was also found in the right associative auditory cortex (Fig. 5c). These results suggest the existence of two levels of perceptual representations, one nonspecific, the other specific to pre-existing spoken words. Distinct neuronal mechanisms would appear to support repetition priming according to the level of representation coded by neurons.

Word-nonspecific and -specific representations

Significant effects of repetition for both words and pseudowords (repetition suppression and enhancement, respectively) were found in two regions of the ventral speech processing pathway (Hickok and Poeppel, 2007), in the right HG extending to the middle STG and in the left PT extending to the posterior STS, suggesting that these regions extract and store physical properties common to both item types. Zatorre et al. (2002) have suggested that the left and right auditory cortices differentially process temporal and spectral auditory information. The left PT is involved in the phonetic analysis of fast formant transitions (Jäncke et al., 2002), whereas the right middle STG regions tend to process spectral variations (Zatorre and Belin, 2001). The PT has been assumed to act as a computational hub (Griffiths and Warren, 2002), where spectro-temporal "templates" are extracted before contacting abstract or "normalized" phonological representations in the later stages of the ventral auditory stream, such as the left posterior STS (Binder et al., 2000). This region may act as an interface between perception and long-term lexical representation of words (Wise et al., 2001), which have been extensively documented as being mediated by the left posterior MTG (Hickok and Poeppel, 2007). Accordingly, we recorded enhanced activity for words relative to pseudowords and word-specific repetition suppression in this region.

Word-specific repetition suppression was also observed in the right anterior STG, which has been linked to talker representations that are independent of the linguistic content (Belin et al., 2004). The formation of spoken word representations specific to a particular voice may rely on a link between word form and talker representation (Schacter et al., 1995b, 2004), preferentially mediated by the right hemisphere (Gonzalez and McLennan, 2007). Repetition suppression in right voice-selective areas may occur with regard to this tight coupling with word-specific regions of the right STS.

Interestingly, in one of these regions (the right middle/posterior STS), we found that HR latency was linked to the word priming RT score and peaked earlier for primed than for unprimed words. This suggests that spoken word priming at least pertains to earlier or shorter neuronal responses in this region. Our finding shows that the size or earliness of the temporal window in which auditory information is integrated may be an important functional property of right middle/posterior STS neurons. This would make sense, given that auditory information over longer timescales, such as syllables, is needed to distinguish between words and pseudowords and may be the word ecological code (Grossberg et al., 1997; Sumner and Samuel, 2007). In line with this idea, the right hemisphere integrates auditory information over longer timescales than the left hemisphere (Boemio et al., 2005; Giraud et al., 2007), probably matching with syllabic variations (Luo and Poeppel, 2007). A similar hypothesis in the visual domain suggests that word recognition relies on the neuronal coding of increasingly large fragments according to a hierarchical progression within the ventral stream (Dehaene et al., 2005), precluding the activation of neuronal units coding sets of letters close to whole words (Vinckier et al., 2007).

Furthermore, our results showed that repetition suppression

Table 1. Peak coordinates of the regions showing a stimulus type by repetition interaction [(words$_{unprimed}$ − words$_{primed}$) − (pseudowords$_{unprimed}$ − pseudowords$_{primed}$)]

Anatomical description	Number of voxels	MNI coordinates			T	p (FDR corrected)
		x	y	z		
Left planum temporale	435	−60	−36	10	4.59	0.024
Left posterior superior temporal sulcus						
Left posterior middle temporal gyrus						
Left precentral gyrus	25	−52	6	30	4.56	0.024
Left inferior frontal gyrus						
Right Heschl gyrus	16	50	−18	10	3.47	0.034
Right superior temporal gyrus						
Left cerebellum	8	−4	−58	−16	3.43	0.035
Right anterior superior temporal gyrus	35	56	0	−8	3.33	0.036
Right superior temporal pole						
Right middle/posterior superior temporal sulcus	12	52	−32	8	3.23	0.037
Left superior temporal pole	20	−52	8	−14	3.19	0.037
Right middle superior temporal sulcus	7	64	−18	4	2.96	0.042

in this right middle/posterior STS was specific to words and correlated with word priming RTs. These findings suggest that spoken word representations are engraved in this part of the auditory cortex by repeated exposure throughout the individual's lifetime. The existence of language-specific memory traces has been proposed previously, but concerned shorter phonetic variations (Näätänen et al., 2001). Our findings also raise the question of the neuronal selectivity to spoken words, an issue still under debate (Cohen and Dehaene, 2004; Price et al., 2005) that cannot be address here. The involvement of a region close to ours in environmental sound priming (Bergerbest et al., 2004) suggests that the functional specialization shared by the various neuronal populations composing this part of the auditory cortex is related to the nature of the process involved, i.e., the encoding of auditory information over long timescales.

Neuronal mechanisms of long-term priming

We showed that perceptual memory traces of spoken words engage various auditory neuronal populations. As discussed above, these distinct neuronal populations appear to process different levels of a word's physical structure, according to their functional properties. It would be interesting to determine whether the same neuronal priming mechanisms are involved in these distinct levels of representation. Regarding word-nonspecific regions (i.e., those showing both word repetition suppression and pseudoword repetition enhancement), none demonstrated a repetition latency shift, which is inconsistent with the facilitation/accumulation model (Henson and Rugg, 2003; James and Gauthier, 2006). It should, however, be noted that repetition latency shifts have been observed in other regions, attesting to the sensitivity of our design to potential shifts in the onset or duration of neural firing related to repetition. One possible alternative, the sharpening model (Desimone, 1996; Wiggs and Martin, 1998), was proposed with regard to repetition suppression but does not specify the nature of the neuronal mechanisms involved in repetition enhancement for pseudowords. Repetition enhancement suggests that single presentations of pseudowords are sufficient to form new cell assemblies (Henson, 2000; Fiebach et al., 2005) that might only be reactivated by primed and not unprimed pseudowords (Henson, 2000, 2003). The neuronal mechanisms by which these new representations are engaged, however, are not yet fully understood. On the one hand, the neuronal encoding of pseudowords may represent key features only. Repetition suppression may arise because neurons coding key features trigger the widespread inhibition of other neurons coding non-key fea-
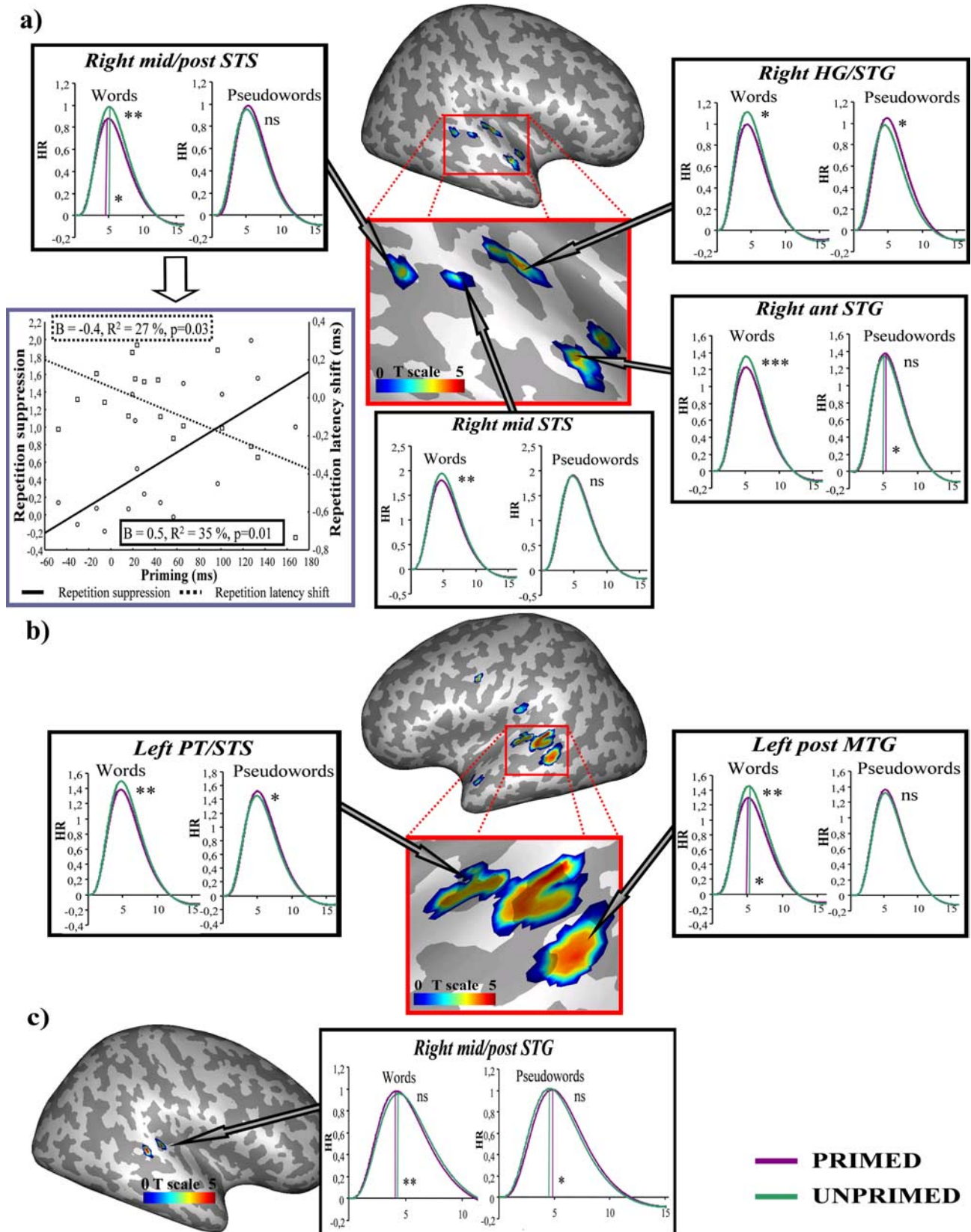
Figure 5. *a–c*, Stimulus type by repetition interaction of the right (*a*) and left (*b*) hemispheres for the activation (voxelwise threshold at $p < 0.05$, FDR corrected for multiple comparisons) and latency (*c*; voxelwise threshold at $p < 0.001$, uncorrected for multiple comparisons) analyses, rendered onto an inflated reconstruction of an MNI template brain. The plots represent the estimated event-related responses based on the sum of the canonical hemodynamic response function and its time derivative, weighted by their parameter estimates. Significant (*$p < 0.05$; **$p < 0.01$) and nonsignificant (ns) differences between mean hemodynamic response amplitude and latency of the primed and unprimed items are shown. The graph (blue/gray frame) represents the relationship between the magnitude of repetition suppression (solid line) or latency shift (dashed line) and behavioral priming (reaction times) in the right middle/posterior superior temporal sulcus. Note that mean latency differences (repetition latency shift) were computed by subtracting unprimed from primed latency maps for each item type. Consequently, negative latency differences in the graph correspond to an earlier HR peak for the primed words.

tures via lateral connections, resulting in a decrease in the fMRI signal pooled over millions of neurons. However, neurons encoding key stimulus features may be more strongly reactivated by repetition (Fize et al., 2000; Grill-Spector et al., 2006). Repetition enhancement would then occur only if those neurons were engaged during the encoding and reactivation of pseudoword representations. However, the formation of new cell assemblies for unfamiliar stimuli may become increasingly elaborate and exhaustive through repeated exposure. Interestingly, consistent with this postulate, a delayed HR without repetition enhancement was found in our study for primed vs unprimed pseudowords in the right middle/posterior and anterior STG. This result may reflect additional neuronal processing during the formation of new representations for pseudowords.

Regarding word-specific regions, the sharpening model, as originally proposed, is insufficient to account for repetition latency shift associated with repetition suppression. However, the co-occurrence of smaller amplitudes and earlier HR peaks prevented us of from ascertaining whether repetition entails briefer neuronal activity, in accordance with the facilitation/accumulation model (Henson and Rugg, 2003; James and Gauthier, 2006), or an earlier decrease in neuronal activity (Henson et al., 2002). Just such reduction of both the activation strength and the onset of neuronal activity with repetition priming was observed by Noguchi et al. (2004), although their demonstration was based on a short-term neural adaptation paradigm. Interpreting our findings in the light of this alternative approach, lower and earlier HR with priming could be an inherent consequence of the process, leading to sharpened representations within the neuronal populations coding high-level representations specific to words: the time taken to reach a sufficient level of inhibition of noncritical neurons by neurons coding key features could be accelerated because of previous exposure (Grill-Spector et al., 2006).

Summary

Using the long-term repetition priming paradigm, we have, for the first time, demonstrated the existence of long-term memory traces coding physical properties of spoken words within the human auditory cortex, paralleling previous results in the visual domain. Auditory word structure appears to be encoded at distinct perceptual levels, either nonspecific or specific to pre-existing word representations. Latency analyses of the HR suggest that the auditory integration timescale and the neuronal mechanisms of priming differ according to these distinct perceptual levels. A classic sharpening model may prevail in auditory regions processing auditory information common to words and pseudowords over short timescales. However, a more integrated model based on a complex interaction between the activation strength and temporal integration of neuronal activity may apply to neuronal populations coding word-specific information within longer temporal windows.

References

Badgaiyan RD, Schacter DL, Alpert NM (1999) Auditory priming within and across modalities: evidence from positron emission tomography. J Cogn Neurosci 11:337–348.

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

Belin P, Fecteau S, Bédard C (2004) Thinking the voice: neural correlates of voice perception. Trends Cogn Sci 8:129–135.

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

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and non-speech sounds. Cereb Cortex 10:512–528.

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

Carlesimo GA, Turriziani P, Paulesu E, Gorini A, Caltagirone C, Fazio F, Perani D (2004) Brain activity during intra- and cross-modal priming: new empirical data and review of the literature. Neuropsychologia 42:14–24.

Church BA, Schacter DL (1994) Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. J Exp Psychol Learn Mem Cogn 20:521–533.

Cohen L, Dehaene S (2004) Specialization within the ventral stream: the case for the visual word form area. NeuroImage 22:466–476.

Dehaene S, Cohen L, Sigman M, Vinckier F (2005) The neural code for written words: a proposal. Trends Cogn Sci 9:335–341.

Desimone R (1996) Neural mechanisms for visual memory and their role in attention. Proc Natl Acad Sci USA 93:13494–13499.

Eger E, Schweinberger SR, Dolan RJ, Henson RN (2005) Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence. NeuroImage 26:1128–1139.

Fiebach CJ, Gruber T, Supp GG (2005) Neuronal mechanisms of repetition priming in occipitotemporal cortex: spatiotemporal evidence from functional magnetic resonance imaging and electroencephalography. J Neurosci 25:3414–3422.

Fize D, Boulanouar K, Chatel Y, Ranjeva JP, Fabre-Thorpe M, Thorpe S (2000) Brain areas involved in rapid categorization of natural images: an event-related fMRI study. NeuroImage 11:634–643.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. NeuroImage 7:30–40.

Friston KJ, Zarahn E, Josephs O, Henson RN, Dale AM (1999) Stochastic designs in event-related fMRI. NeuroImage 10:607–619.

Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J (2002) Classical and Bayesian inference in neuroimaging: applications. NeuroImage 16:484–512.

Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15:870–878.

Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiack RSJ, Laufs H (2007) Endogenous cortical rhythms determine cerebral specialization for speech perception and production. Neuron 56:1127–1134.

Gonzalez J, McLennan CT (2007) Hemispheric differences in indexical specificity effects in spoken word recognition. J Exp Psychol Hum Percept Perform 33:410–424.

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 10:14–25.

Grossberg S, Boardman I, Cohen C (1997) Neural dynamic of variable-rate speech categorization. J of Exp Psychol Hum Percept Perform 23:418–503.

Habeck C, Hilton HJ, Zarahn E, Brown T, Stern Y (2006) an event-related fMRI study of the neural network underlying repetition suppression and reaction time priming in implicit visual memory. Brain Res 1075:133–141.

Henson RN (2000) Neuroimaging evidence for dissociable forms of repetition priming. Science 287:1269–1272.

Henson RN (2003) Neuroimaging studies of priming. Prog Neurobiol 70:53–81.

Henson RN, Rugg MD (2003) Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia 41:263–270.

Henson RN, Price CJ, Rugg MD, Turner R, Friston KJ (2002) Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. NeuroImage 15:83–97.

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–401.

James TW, Gauthier I (2006) Repetition-induced changes in BOLD re-

sponse reflect accumulation of neural activity. Hum Brain Mapp 27:37–46.

Jäncke L, Wüstenberg T, Scheich H, Heinze HJ (2002) Phonetic perception and the temporal cortex. NeuroImage 15:733–746.

Lebreton K, Desgranges B, Landeau B, Baron JC, Eustache F (2001) Visual priming within and across symbolic format using a tachistoscopic picture identification task: a PET study. J Cogn Neurosci 13:670–686.

Luce PA, Lyons EA (1998) Specificity of memory representations for spoken words. Mem Cogn 26:708–715.

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

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

Noguchi Y, Inui K, Kakigi R (2004) Temporal dynamics of neural adaptation effect in the human visual ventral stream. J Neurosci 24:6283–6290.

Orfanidou E, Merslen-Wilson WD, Davis MH (2006) Neural response suppression predicts repetition priming of spoken words and pseudowords. J Cogn Neurosci 18:1237–1252.

Price C, Thierry G, Griffiths T (2005) Speech-specific auditory processing: where is it? Trends Cogn Sci 9:271–276.

Schacter DL, Church BA (1992) Auditory priming: implicit and explicit memory for words and voices. J Exp Psycho Learn Mem Cogn 18:915–930.

Schacter DL, Reiman E, Uecker A, Polster MR, Yun LS, Cooper LA (1995a) Brain regions associated with retrieval of structurally coherent visual information. Nature 376:587–590.

Schacter DL, Church BA, Bolton E (1995b) Implicit memory in amnesic patients: impairment of voice-specific priming. Psychol Sci 6:20–25.

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

Schott BH, Richardson-Klavhen A, Henson RN, Becker C, Heinze HJ, Düzel E (2006) Neuroanatomical dissociation of encoding processes related to priming and explicit memory. J Neurosci 26:792–800.

Sumner M, Samuel AG (2007) Lexical inhibition and sublexical facilitation are surprisingly long lasting. J Exp Psychol Learn Mem Cogn 33:769–790.

Swick D, Miller KM, Larsen J (2004) Auditory repetition priming is impaired in pure alexic patients. Brain Lang 89:543–553.

Thiel CM, Henson RN, Dolan RJ (2002) Scopolamine but not lorazepam modulates face repetition priming: a psychopharmacological fMRI study. Neuropsychopharmacology 27:282–292.

Tulving E, Schacter DL (1990) Priming and human memory systems. Science 247:301–306.

Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L (2007) Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. Neuron 55:143–156.

Vuilleumier P, Henson RN, Driver J, Dolan RJ (2002) Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nat Neurosci 5:491–499.

Wager TD, Nichols TE (2003) Optimization of experimental design in fMRI: a general framework using a genetic algorithm. NeuroImage 18:293–309.

Westbury CF, Zatorre RJ, Evans AC (1999) Quantifying variability in the planum temporale: a probability map. Cereb Cortex 9:392–405.

Wiggs CL, Martin A (1998) Properties and mechanisms of perceptual priming. Curr Opin Neurobiol 8:227–233.

Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA (2001) Separate neural subsystems within "Wernicke's area." Brain 124:83–95.

Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. Cereb Cortex 11:946–953.

Zatorre RJ, Belin P, Penhume V (2002) Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6:37–46.