

Thalamic and parietal brain morphology predicts auditory category learning



Mathias Scharinger^{a,*}, Molly J. Henry^a, Julia Erb^a, Lars Meyer^b, Jonas Obleser^a

^a Max Planck Research Group "Auditory Cognition", Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

^b Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

Article history:

Received 28 May 2013

Received in revised form

2 September 2013

Accepted 4 September 2013

Available online 13 September 2013

Keywords:

Auditory categorization

Acoustic cues

Cue utilization

Voxel-based morphometry

Auditory thalamus

Parietal cortex

ABSTRACT

Auditory categorization is a vital skill involving the attribution of meaning to acoustic events, engaging domain-specific (i.e., auditory) as well as domain-general (e.g., executive) brain networks. A listener's ability to categorize novel acoustic stimuli should therefore depend on both, with the domain-general network being particularly relevant for adaptively changing listening strategies and directing attention to relevant acoustic cues. Here we assessed adaptive listening behavior, using complex acoustic stimuli with an initially salient (but later degraded) spectral cue and a secondary, duration cue that remained nondegraded. We employed voxel-based morphometry (VBM) to identify cortical and subcortical brain structures whose individual neuroanatomy predicted task performance and the ability to optimally switch to making use of temporal cues after spectral degradation. Behavioral listening strategies were assessed by logistic regression and revealed mainly strategy switches in the expected direction, with considerable individual differences. Gray-matter probability in the left inferior parietal lobule (BA 40) and left precentral gyrus was predictive of "optimal" strategy switch, while gray-matter probability in thalamic areas, comprising the medial geniculate body, co-varied with overall performance. Taken together, our findings suggest that successful auditory categorization relies on domain-specific neural circuits in the ascending auditory pathway, while adaptive listening behavior depends more on brain structure in parietal cortex, enabling the (re)direction of attention to salient stimulus properties.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The categorization of novel sounds reflects the ability to assign behaviorally relevant meaning to acoustic information. Recent research has contributed to a better understanding of the neural bases of both non-speech (Griffiths, 2001; Griffiths, Uppenkamp, Johnsrude, Josephs, & Patterson, 2001; Husain et al., 2006; Sharda & Singh, 2012) and speech categorization (Binder, Liebenthal, Possing, Medler, & Ward, 2004; Blumstein, Myers, & Rissman, 2005; Guenther, Nieto-Castanon, Ghosh, & Tourville, 2004; Hickok & Poeppel, 2007; Liebenthal et al., 2010; Myers, Blumstein, Walsh, & Eliassen, 2009; Scott & Johnsrude, 2003). In particular, it has been shown that detailed spectral and temporal processing is subserved by bilateral temporal areas (i.e. domain-specific auditory regions), while more categorical aspects of processing additionally involve parietal areas (i.e. domain-general executive regions). However, relatively little is known regarding the extent to which morphology in these areas predicts adaptive behavior

during categorization, for instance, when a once informative acoustic cue is rendered less informative. The switch to an alternative acoustic cue in such a situation can be referred to as *cue switching* (also called *cue weighting*, cf. Francis, Kaganovich, & Driscoll-Huber, 2008; Holt & Lotto, 2006; Lipski, Escudero, & Benders, 2012). The question we ask here is: is this ability directly supported by brain areas that process acoustic cues (i.e., temporal auditory areas), or does cue switching require the deployment of attentional resources (i.e., from executive areas)?

In the current study, we examined the extent to which brain morphology predicts successful categorization of novel acoustic stimuli or (optimal) cue switching, necessitated by the degradation of a once informative acoustic cue. We hypothesized that successful categorization should partially depend on temporal auditory areas that support the processing of spectral and temporal cues, but should also rely on parietal areas that support categorical processing and allow for the adaptive re-direction of attentional resources towards the most informative acoustic cue.

From a behavioral point of view, cue switching seems to critically depend on the ability to rapidly re-direct attention to the most salient cue in a given listening situation (Francis, Baldwin, & Nusbaum, 2000; Francis & Nusbaum, 2002; Francis et al., 2008). This can be achieved via explicit task instructions

* Correspondence to: Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1A, 04103 Leipzig, Germany. Tel.: +49 341 9940 2482.

E-mail address: mscharinger@cbs.mpg.de (M. Scharinger).

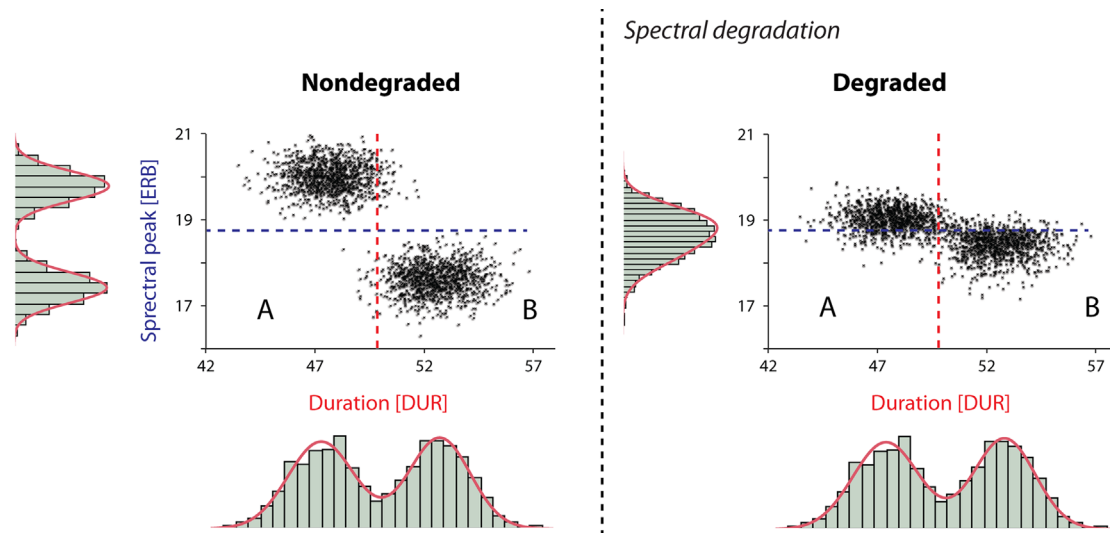


Fig. 1. Stimulus characteristics in two-dimensional space of spectral peak (ERB; y-axis) and log 10 duration (DUR; x-axis). Note that tones in the nondegraded condition (left) show a bimodal spectral peak distribution, with clear separation between categories A and B. Degraded tones (right) have a smeared (i.e., unimodal) spectral peak distribution. By contrast, bimodal duration distributions are not affected by degradation. For this reason, the primary cue for categorization in the nondegraded condition of the experiment was spectral peak, while it was duration in the degraded condition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(top-down; Wang & Humes, 2008) or as a result of changes in the stimulus material (bottom-up; Holt & Lotto, 2006); the latter effect is sometimes referred to as ‘attentional capture’ in analogy to vision research (Yantis, 1993). The involvement of brain networks supporting attentional aspects in auditory categorization (Gazzaley & Nobre, 2012; Raz & Buhle, 2006) has thus far received only little attention. Instead, most structural and functional brain imaging studies have focused on areas in or around primary and secondary auditory cortices (superior temporal gyrus [STG], Heschl's gyrus [HG], planum temporale [PT], e.g. Bermudez, Lerch, Evans, & Zatorre, 2009; Hall et al., 2002; Sharda & Singh, 2012). This also holds for speech processing (e.g. Hutchison, Blumstein, & Myers, 2008), but there is accumulating evidence that categorical aspects of processing additionally involve frontal and parietal areas (Binder et al., 2004; Blumstein et al., 2005; Guenther et al., 2004; Hickok & Poeppel, 2007; Liebenthal et al., 2010; Myers et al., 2009; Scott & Johnsrude, 2003) as well as subcortical networks (Frith & Friston, 1996; von Kriegstein, Patterson, & Griffiths, 2008; Tang, Yang, & Suga, 2012).

Structural studies using voxel-based morphometry (VBM) commonly relate the probability¹ of gray matter in temporal brain areas to particular auditory skills, that is, in musicians compared to non-musicians (Bendor & Wang, 2005; Bermudez & Zatorre, 2005; Gaser & Schlaug, 2003; Schneider et al., 2002; Zatorre, Belin, & Penhune, 2002). Such correlations between musical skill and gray-matter probability are however not restricted to domain-specific (auditory) areas, as for instance evidenced by Foster and Zatorre (2010). They identified the intraparietal sulcus as additional area where gray-matter probability correlated with musical aptitude. The fact that auditory skills co-vary with brain morphology not only in domain-specific auditory areas, but also in rather domain-general areas, seems to hold for speech as well (Golestani, Price, & Scott, 2011). The authors found that the size of the left pars opercularis positively correlated with the degree of phonetic training in their participants.

In this regard, functional brain imaging studies have shown sound categorization to involve parts of the domain-general attention network (Anderson, Ferguson, Lopez-Larson, & Yurgelun-Todd, 2010; Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000), including bilateral middle temporal and supra-marginal gyri (Guenther et al., 2004), inferior frontal gyri, and (inferior) parietal lobules (Husain et al., 2006). In speech, the involvement of these areas has been interpreted against the background of the goodness-of-fit between perceptual evidence and category information, and has additionally been related to the segmentation of the acoustic signal into meaningful units (Burton, Small, & Blumstein, 2000; Blumstein et al., 2005; Friederici, 2011; Myers, 2007; Myers & Swan, 2012). However, it remains to be seen whether increased gray-matter probability in parietal areas may correspond to a better ability to allocate attentional resources to relevant acoustic information during auditory categorization. For this reason, in the current study, we adapted an auditory-categorization design previously applied by Goudbeek, Swingley, and Smits (2009) that used complex tones, varying in their spectral peaks and durations. We modified this design (i.e., the statistical distribution of stimulus tokens) such that spectral peak was the primary cue for categorization in the first condition of the experiment, with duration constituting a secondary cue (Fig. 1, left). Halfway through the experiment, spectral cues were degraded, and the duration cue, unaffected by degradation, became relatively more informative (Fig. 1, right).

Tendencies of using spectral (spectral peak) vs. temporal (duration) cues for categorization were assessed by means of logistic regressions. Logistic regressions can be used to predict category membership decisions on the basis of stimulus properties (Hosmer & Lemeshow, 2000). We further used VBM to identify cortical and subcortical brain structures where individual differences in brain morphology (here: gray-matter probability) would predict task performance and the ability to optimally switch from primarily using spectral cues in the beginning of the experiment to primarily using temporal cues after spectral degradation.

We expected that auditory categorization would depend on individual abilities related to (1) processing spectro-temporal cues and (2) directing attention to the most informative of these cues in the two experimental conditions (nondegraded vs. degraded spectral cues). For this reason, we hypothesized finding significant

¹ Note that we chose to refer to gray-matter volume and density (Mechelli, Price, Friston, & Ashburner, 2005) as gray-matter probability, on the basis of the VBM analyses that identify particular voxels as gray-matter voxels.

correlations (1) between gray-matter probabilities in auditory temporal and/or subcortical regions and overall perceptual sensitivity in auditory categorization and (2) between gray-matter probabilities in parietal areas and optimal cue utilization (i.e. spectral cues in the nondegraded, and temporal cues in the degraded condition).

2. Materials and methods

2.1. Participants

Thirty-six healthy volunteers, recruited from the participant database of the Max Planck Institute for Human Cognitive and Brain Sciences (all right-handed, 18 female, 18 male, age range 20–31 years, mean age 25.7; standard deviation [SD] = 2.8 years) participated in the experiment. All were native speakers of German with no self-reported hearing impairments or neurological disorders. Participants gave written informed consent and received financial compensation for their participation. All procedures followed the guidelines of the local ethics committee (University of Leipzig) and were in accordance with the Declaration of Helsinki.

2.2. Stimuli

Stimuli were complex tones varying in spectral filtering and duration. They were designed according to a previously published protocol (Goudbeek et al., 2009). In short, stimulus creation involved the modification of a base signal that consisted of 17 sinusoidal components. The lowest component frequency was 500 Hz, and the frequency ratio between successive components equalled 0.67. Components were summed without weighting. Modulation of the spectral peak was achieved by applying a second-order infinite-impulse response filter with a bandwidth of 0.2 times its target frequency.

Categories (A and B) were defined by the dimensions *spectral peak* (target frequency of filter) and *duration*. Stimulus distributions, stretched along the

duration axis (Fig. 1), were formed by generating separate bivariate normal distributions with equal standard deviations, σ , and means, μ , at 40 equally-spaced locations along the duration dimension. Twenty-five tokens were retained for each of the 40 distributions, resulting in a total of 1000 stimuli per distribution (Table 1; Fig. 2). Waveforms and spectrograms of exemplary nondegraded and degraded stimuli are shown in Fig. 2.

The spread of filter frequencies (i.e. standard deviation, σ , in the frequency dimension) was such that the two categories clearly differed (did not overlap) while the distributions of durations were arranged such that 1/3 of the exemplars in category A and B overlapped in duration (Fig. 1A, left panel). Spectral peak and duration manipulations were according to normed physical scales that have been previously shown to afford psychophysical comparability of the two dimensions (Smits, Sereno, & Jongman, 2006). Specifically, frequencies were converted to the equivalent rectangular bandwidth (ERB) scale that approximates the bandwidths of the auditory filters in human hearing (Glasberg & Moore, 1990), and durations were converted to a log 10 scale (DUR; cf. Smits et al., 2006).

For the intended salience manipulation, that is, degradation of spectral cues, degraded versions of tones were generated using noise vocoding (Drullman, Festen, & Plomp, 1994; Shannon, Zeng, Kamath, Wyganski, & Ekelid, 1995). In short, this procedure involves dividing the raw signal into an arbitrary number of frequency bands (here: four; e.g., Erb, Henry, Eisner, & Obleser, 2012; Shannon et al., 1995), extracting the amplitude envelope from each band and reapplying it to bandpass-filtered noise carriers, thereby smearing spectral details. Noise-vocoding was applied to all stimuli, irrespective of category, using MATLAB 7.9 as described in Rosen, Faulkner, and Wilkinson (1999), with settings as reported in Erb et al. (2012). Degraded stimuli maintained their amplitude envelope features and, most importantly, original duration; however, degraded stimuli showed a change in the location and spread of spectral peak (Table 1; Fig. 1, right, Fig. 2, bottom).

All stimuli were normalized to equal root-mean-square (RMS) intensity and presented at ~60 dB sound-pressure level (SPL). Linear onset and offset ramps (5 ms) ensured that acoustic artifacts were minimized.

2.3. Experimental procedure

First, a short practice block with 10 tones from each category (A and B) familiarized participants with the task. During the main experiment, exemplars were presented from categories A and B randomly with equal probability. On each trial, a single stimulus exemplar was presented. Then, a visual response prompt was presented (a green light appeared on a screen), and participants indicated whether the exemplar belonged to category A or category B by pressing one of two keys on a button box; button assignment was counterbalanced across participants. Following the response, participants received corrective feedback ('Correct'/'Incorrect'), which was displayed for 1 s in the middle of the screen. This experimental design was previously used to investigate supervised learning of novel acoustic stimuli (Scharinger, Henry, & Obleser, 2013).

Four blocks containing 60 trials each were presented, with short breaks between blocks. In the first two blocks, stimuli were nondegraded (Fig. 1A, left

Table 1
Means of spectral peak (in ERB) and duration (DUR; log 10 duration) for stimulus categories A and B in the nondegraded and degraded conditions. Standard deviations (σ) are given in parentheses.

Stimulus category	Nondegraded		Degraded	
	A	B	A	B
Spectral peak [ERB]	20.00 (0.31)	17.00 (0.31)	16.80 (0.31)	15.50 (0.31)
Duration [DUR]	47.70 (1.31)	52.53 (1.31)	47.70 (1.31)	52.53 (1.31)

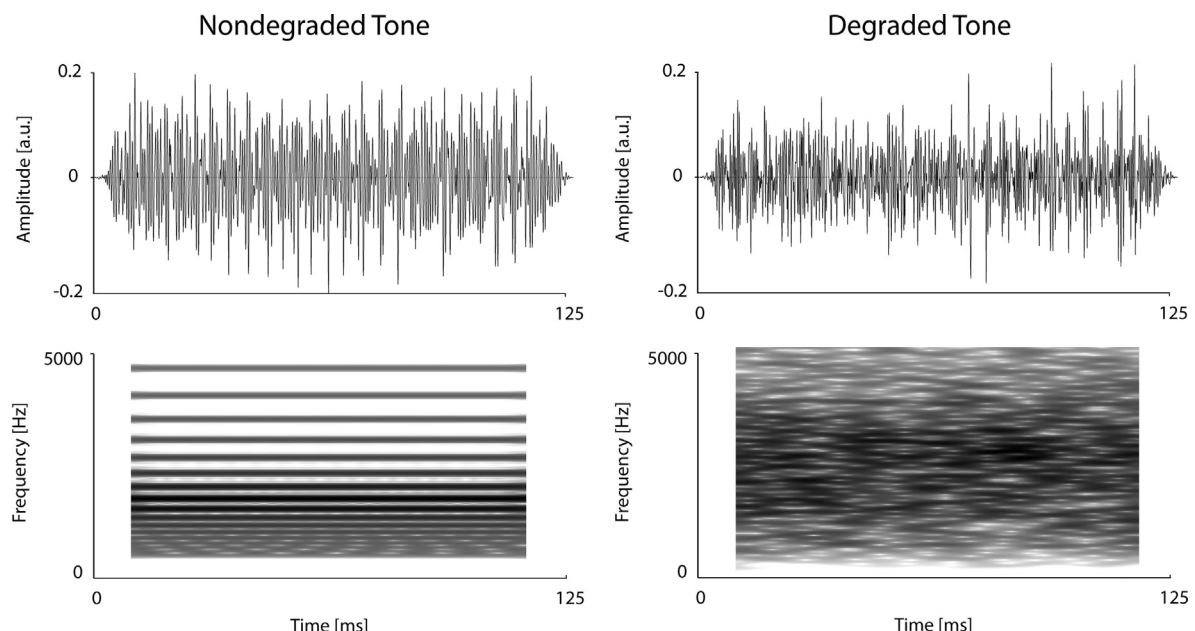


Fig. 2. Acoustic analyses of tone stimuli. The left panel illustrates wave form (top) and spectrogram of a nondegraded tone, while the right panel shows waveform and spectrogram of a degraded tone. Spectrograms were calculated using a fast-Fourier analysis with a Hanning-window (size: 20 ms, step size: 2 ms).

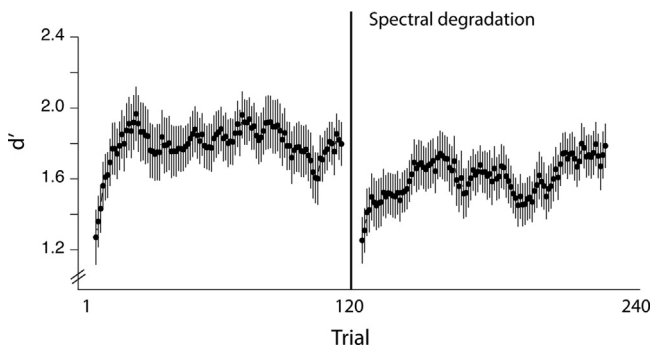


Fig. 3. Perceptual sensitivity (d') obtained from sliding windows over nondegraded and degraded trials per participant (window size=20 trials, step size=1 trial). Note that learning, approximated by fast changes in perceptual sensitivity in the beginning, occurred relatively fast. Spectral degradation caused temporary worse performance, but recovery took place immediately, even though the level of previous performance (nondegraded condition) was not reached any more.

panel), while in the second two blocks, stimuli were degraded (Fig. 1A, right panel). The duration of the entire experiment was 50 min.

2.4. Analyses of behavioral data

We used three behavioral measures as regressors of interest in the VBM analyses: one for overall performance, and two indexing cue utilization.

First, overall performance was assessed by d' , a measure of perceptual sensitivity that is independent of response bias. Perceptual sensitivity, d' , was calculated from proportions of hits and false alarms (Macmillan & Creelman, 2005), where hits were defined as proportions of “category-A” responses to category-A stimuli, and false alarms were defined as “category-A” responses to “category-B” stimuli. In order to get more dynamic estimates of category learning, we also calculated d' values in sliding windows across all items. To that end, we calculated d' values for hit and false alarm averages in windows consisting of 20 items, and moved these windows using a step size of 1 item. This effectively yields finer-grained d' values, mimicking single-trial measures. The pattern of d' values obtained from the sliding windows is illustrated in Fig. 3.

Second, cue utilization was assessed by logistic regressions, where the likelihood of a category-A response was predicted from spectral peak and duration. Absolute β values for spectral peak and duration could then be compared, with higher values reflecting a stronger influence of the relevant dimension for the categorical response (Goudbeek et al., 2009; Scharinger et al., 2013). To that end, we defined the β index as a measure of cue utilization based on the difference between β values for spectral peak and duration, in relation to the sum of both β values:

$$\beta \text{ index} = \frac{\beta_{\text{Duration}} - \beta_{\text{Frequency}}}{\beta_{\text{Duration}} + \beta_{\text{Frequency}}} \quad (1)$$

According to (1), a positive β index reflects the tendency to use temporal cues more than spectral cues, while a negative β index reflects the tendency to use spectral cues more than temporal cues.

Third, we reasoned that an optimal cue switch would be characterized by more negative β indices in the first condition of the experiment (primary cue: spectral peak) than in the second condition of the experiment (primary cue: duration). This is expressed by the difference of the β indices:

$$\text{switch index} = \beta \text{ index}_{\text{Degraded}} - \beta \text{ index}_{\text{Nondegraded}} \quad (2)$$

where an optimal cue switch is reflected by a positive switch index (more reliance on spectral cues in the beginning of the experiment than after degradation), while a non-optimal cue switch is reflected by a negative switch index (more reliance on temporal cues in the beginning of the experiment than after degradation).

2.5. Structural image acquisition and analysis

Structural images were T1-weighted anatomical scans acquired with a 12-channel head coil using an MPRAGE sequence (repetition time=1300 ms, echo time=3.46 ms, flip angle=10°, 176 sagittal slices, acquisition matrix of 256 × 240, voxel size=1 mm³) on a 3 T Siemens TIM Trio scanner. Structural scans of all 36 participants had been acquired, on average, 29 months prior to the experiment (SD=18 months), and time since image acquisition was used as a regressor of no interest in all analyses reported (see below).

Structural images were subjected to voxel-based morphometry (VBM), a technique implemented in SPM8 (Wellcome Trust, Department of Imaging Neuroscience, London, UK) that allows assessment of correlations in regional gray-matter probability with explanatory variables (Ashburner & Friston, 2000). The

focus on gray matter was motivated by our interest in local computational efficacy of particular brain regions, rather than long-distance connectivity (cf. Ashburner & Friston, 2000). First, raw images were segmented into gray matter, white matter, and cerebro-spinal fluid images. Gray- and white-matter images were further processed using the DARTEL algorithm that generates group-specific gray- and white-matter templates (Ashburner, 2007). Spatial deformations were used to map the individual data onto their group-specific template. Subsequently, the group template was normalized to MNI space. Images were then smoothed with an isotropic Gaussian kernel of 8 mm FWHM and adjusted for the difference in volume before and after normalization. Individual voxels of these images represent a measure of regional volume and density (Ashburner, 2009).

We analyzed the modulated gray-matter images in multiple-regression analyses using general linear models (GLM) with per-participant averages of overall performance (d' values), β index (overall tendency to utilize spectral or temporal cues) and cue switch (optimal switch from utilizing spectral cues to utilizing temporal cues) in separate models.

Gender, age (in years), and time since T1 acquisition (in months) were included as regressors of no interest (Pelle, Cusack, & Henson, 2012). All models used global normalization to correct for total gray-matter volume, since our interest was not in absolute, but rather relative, local gray-matter volume.

Results were corrected for a whole-brain type-I error of less than 0.05, based on cluster extent ($k=50$; $p < 0.001$ [voxel-level]). However, cluster extent corrections can be problematic in VBM studies due to the potential non-stationary smoothness of the data. Hence, we used a non-stationary cluster extent correction (Hayasaka, Phan, Liberzon, Worsley, & Nichols, 2004) within the VBM8 toolbox, effectively correcting the cluster extent threshold for non-isotropic smoothness ($p < 0.001$ [voxel-level]; available at <http://dbm.neuro.uni-jena.de/vbm/download/>).

2.6. Statistical analyses

For the behavioral data, repeated-measures analyses of covariance (ANCOVAs) were calculated to examine the influence of cue utilization (absolute β values and β index) on categorization performance (d') in the first and second condition of the experiment. Both ANOVAs used d' as dependent variable; the first ANOVA comprised the covariates β spectral and β duration in interaction with condition (referring to the nondegraded and degraded condition of the experiment), while the second used the covariate β index, also in interaction with condition.

Whole-brain regression analyses on gray-matter probability using the three behavioral measures introduced above were carried out in SPM8 (Wellcome Trust, Department of Imaging Neuroscience, London, UK). We also compared correlations between gray-matter probability and our behavioral measures across conditions (nondegraded, degraded).

3. Results

3.1. Performance measures show successful categorization of novel acoustic stimuli

Overall, participants' accuracy was above chance level (on average, 87% correct; $t(35)=19.14$, $p < 0.001$; average $d'=1.58$, $t(35)=23.12$, $p < 0.001$), reflecting successful categorization of the novel acoustic stimuli. Perceptual sensitivity differed between conditions (nondegraded $d'=1.67$ vs. degraded $d'=1.49$; $t(35)=2.10$, $p < 0.05$; Fig. 4 top). Perceptual sensitivity obtained from the sliding window approach showed the following temporal profile (Fig. 3): participants very rapidly improved their categorization within the first 20–30 trials. After spectral degradation, performance dropped, but quickly recovered, albeit the previous average level of performance was not reached any more (see t -test above).

3.2. Strategy use differs between nondegraded and degraded conditions

Averaged β values were more negative for the nondegraded (-0.41) than the degraded (-0.24) condition ($t(35)=1.90$, p (one-tailed) < 0.05 ; Fig. 4 bottom). Thus, in the nondegraded condition of the experiment, participants relied on spectral cues to a higher degree than in the degraded condition. Note though that there was an apparent resilience to entirely give up on the previously used spectral cues (i.e. the average β index was still negative).

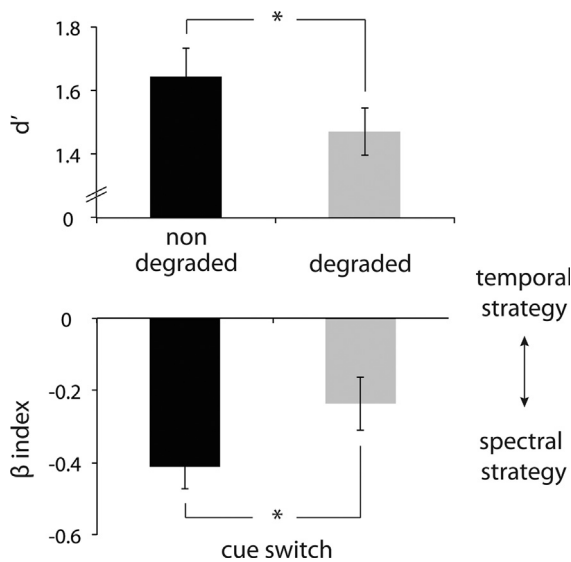


Fig. 4. Performance (as measured from d' ; top) and cue utilization (as measured from β index; bottom) in the nondegraded and degraded conditions of the experiment. Note that a more negative β index reflects more evidence for spectral cue utilization, while a more positive β index reflects more evidence for temporal cue utilization.

3.3. Strategy use predicts performance

The repeated-measures ANOVA on d' as dependent variable, with the independent variable condition (nondegraded; degraded) and the covariates β spectral and β duration, showed main effects of condition ($F(1,35)=10.5$, $p < 0.01$, $\eta_p^2 = 0.14$), β spectral ($F(1,35)=13.6$, $p < 0.001$, $\eta_p^2 = 0.18$), and β duration ($F(1,35)=13.3$, $p < 0.001$, $\eta_p^2 = 0.17$). The main effect condition replicated the performance differences between experimental conditions, while the main effects β spectral and β duration showed that overall utilization of either spectral peak or duration improved performance.

Crucially, spectral cue utilization differentially impacted performance in the two experimental conditions, as seen in the significant interaction for β spectral and condition ($F(1,35)=5.1$, $p < 0.01$, $\eta_p^2 = 0.07$). Temporal cue utilization was independent of condition (β duration $\times$ condition, $F(1,35) < 1$). However, separate analyses for the nondegraded and degraded condition revealed that both β spectral ($F(1,35)=13.3$, $p < 0.001$, $\eta_p^2 = 0.30$) and β duration ($F(1,35)=6.3$, $p < 0.01$, $\eta_p^2 = 0.17$) were significant in the nondegraded condition, while in the degraded condition, only β duration ($F(1,35)=7.8$, $p < 0.01$, $\eta_p^2 = 0.16$), but not β spectral ($F(1,35)=1.9$, $p < 0.18$, $\eta_p^2 = 0.05$) was significant. Thus, both spectral peak and duration cue utilization led to better performance in the nondegraded condition.

In sum, the behavioral data indicate that participants differentially used acoustic cues for categorization, with stronger reliance on spectral cues in the first condition and relatively stronger reliance on temporal cues in the second condition of the experiment. Further, the expected tendency towards using temporal cues under degradation predicted better performance.

3.4. Whole-brain analyses

Multiple-regression analyses on the whole-brain level resulted in a number of significant clusters that are summarized in Table 2. These comprised subcortical, thalamic, as well as parietal areas, and pre- and postcentral gyri. Some of the clusters that are all based on positive correlations with the three behavioral measures, as discussed in more detail below, are further illustrated in Fig. 5.

Table 2

Correlations between gray-matter probability and behavioral performance measure: d' , β index and switch index, corrected for non-stationary cluster extent ($p < 0.001$). Coordinates (x ; y ; z) are given in MNI space [MGB=medial geniculate body].

Area	Coordinates	Z	Extent (K)
<i>Perceptual sensitivity (d')</i>			
Left thalamus	−20; −21; 1	3.35	209
Left thalamus (MGB)	−14; −25; −6	3.27	
Right postcentral gyrus	24; −31; 49	3.97	233
Right postcentral gyrus	15; −39; 64	3.76	68
Right precentral gyrus	58; −18; 43	3.63	172
<i>β Index (more temporal cue use overall)</i>			
Right inferior parietal lobule	43; −40; 45	3.37	55
<i>β Index (more temporal cue use in degraded than in nondegraded condition)</i>			
Left precentral gyrus	−24; −29; 57	3.81	386
Right inferior parietal lobule	40; −37; 39	3.34	36
<i>Switch</i>			
Left precentral gyrus	−24; −28; 58	4.0	510
Right inferior parietal lobule	40; −37; 37	3.49	77

3.4.1. Correlates of performance

Success of overall performance co-varied with gray-matter probability in an extensive left thalamic area, comprising the ventral posterior medial and lateral nucleus and the medial geniculate body (Table 2). The comparisons across conditions did not yield any significant clusters.

Furthermore, performance also positively correlated with gray-matter probability in right postcentral gyrus, extending into somatosensory cortex (Brodmann Area [BA] 3), and, in the second cluster (Table 2), extending into motor cortex (BA 4). Finally, gray-matter probability positively correlated with d' in right precentral gyrus, comprising motor (BA 4) and pre-motor cortices (BA 6). Thus, increased gray-matter probability in subcortical, thalamic areas as well as in post- and precentral gyrus corresponded to better categorization performance as measured in d' .

3.4.2. Correlates of cue utilization

Overall, β index positively correlated with gray-matter probability in right inferior parietal lobule (BA 40). Also, the comparison for the β index across conditions (degraded vs. nondegraded) showed two significant clusters, one in left precentral gyrus, comprising BA 4 and 6, and one in right inferior parietal lobule, comprising supramarginal gyrus and BA 40 (Table 2).

Subsequent post-hoc whole-brain analyses on β indices from the nondegraded and degraded condition of the experiment showed significant correlations only in the degraded condition. Here, more positive values corresponded to increased gray-matter probability in bilateral inferior parietal lobule (left cluster: Extent (K)=226, $Z=4.31$, peak coordinate $x=-38$, $y=-40$, $z=51$; BA 40; right cluster: Extent (K)=565, $Z=4.47$, peak coordinate $x=42$, $y=-39$, $z=40$; BA 40), with the right cluster extending into supra-marginal gyrus.

The switch index correlated positively with gray-matter probability in right inferior parietal lobule (including parts of supra-marginal gyrus and BA 40) and precentral gyrus, extending into parietal lobe and primary motor cortex (BA 4; Table 2).

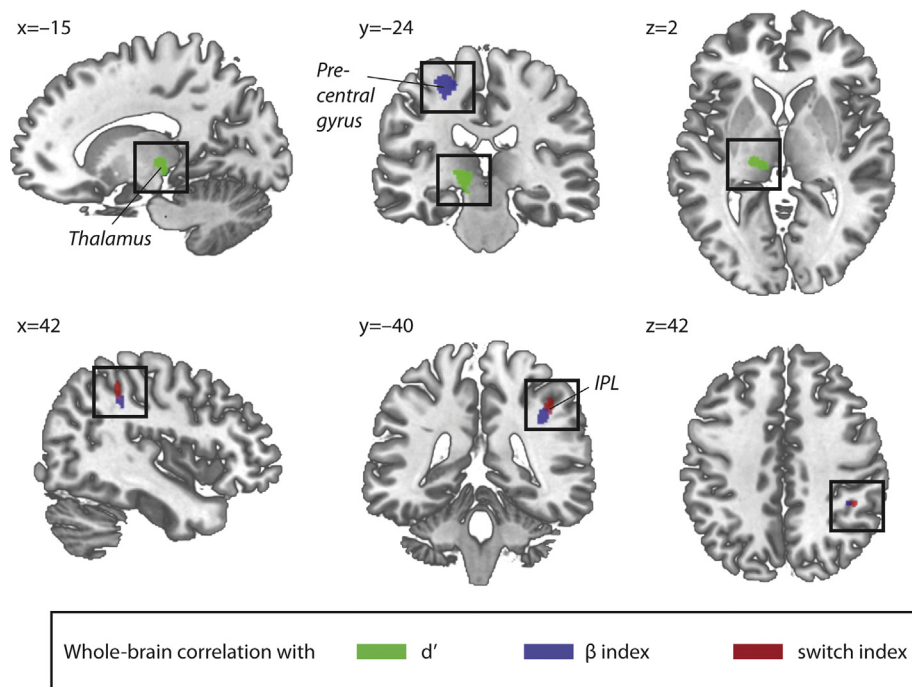


Fig. 5. Significant clusters (corrected for non-stationary cluster extent $p < 0.001$) of whole-brain VBM analyses. Correlations of gray-matter probability with d' (green), β index (blue) and switch index (red) [IPL=Inferior Parietal Lobule]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

In an auditory-categorization experiment, participants learned to assign categories to novel and complex acoustic stimuli (tones) that differed in spectral peak and duration. Initially, the primary cue for categorization was spectral peak (Fig. 1, left), while after spectral degradation in the second half of the experiment, the primary cue was stimulus duration (Fig. 1, right), which was left unaffected by the manipulation.

Behavioral results based on logistic regression confirmed that participants used stimulus cues that were most informative (spectral cues in first condition, temporal cues in second condition). The major novel finding of this study, however, is that subcortical (thalamic) and parietal gray-matter probability predicted performance and cue utilization, respectively, during auditory categorization.

4.1. Inferior parietal morphology predicts cue utilization

The behavioral results of our study suggested that overall performance depended on employing an optimal categorization strategy. According to our stimulus design (Fig. 1), participants should have initially used spectral peak as the primary cue for categorization and then switched to duration in the second half of the experiment, when spectral peak was rendered uninformative by degradation. In fact, the majority of participants (62%) showed this switch. Further, more evidence of temporal cue utilization in the degraded condition (as measured by higher absolute β values) was predictive of better performance.

The VBM correlations with β index and switch index revealed clusters in parietal cortex. More precisely, gray-matter probability in (predominantly right) inferior parietal lobule (IPL) positively correlated with evidence for temporal cue utilization and optimal cue switching. That is, IPL morphology was predictive of whether participants would use duration in order to categorize the novel acoustic stimuli in our experiment after application of spectral degradation.

The inferior parietal lobule as part of the fronto-parietal executive network (Corbetta et al., 2000; Posner & Dehaene, 1994) has been implicated in selective attention (Shaywitz et al., 2001; Salmi, Rinne, Koistinen, Salonen, & Alho, 2009) and attentional control during auditory processing (Hill & Miller, 2010), spatial and non-spatial auditory working memory (Alain & Arnott, 2000; Brunetti et al., 2008; Leung & Alain, 2011; Marshuetz, Smith, Jonides, DeGutis, & Chenevert, 2000; Rinne et al., 2007; Schulze & Koelsch, 2012), auditory imagery (Shergill, Bullmore, Simmons, Murray, & McGuire, 2000; Shergill et al., 2001) and sound (phoneme) categorization (Guenther et al., 2004; Husain et al., 2006; Turkeltaub & Coslett, 2010). Furthermore, an area in close vicinity and also comprising parts of BA 40, the intraparietal sulcus (IPS) has been found to be critically involved in structuring sensory input, perceptual organization and goal-directed behavior (Cusack, 2005; Cohen, 2009; Teki, Chait, Kumar, von Kriegstein, & Griffiths, 2011). Finally, recent findings from our lab point to a specific role of the inferior parietal cortex in directing attention towards or away from particular temporal stimulus features (i.e., duration; Henry, Herrmann, & Obleser, in press).

Overall, these data support the assumption that the IPL maintains incoming acoustic stimuli in short-term memory for comparing them with an emergent category structure representation that participants are acquiring in our experiment. For such comparisons, it is beneficial to direct attention to the most informative acoustic stimulus cue, and segregate it from irrelevant information (e.g. figure-ground segregation, as supported by intraparietal sulcus, Teki et al., 2011). In the current task, this involved a switch to listening more closely to duration in the degraded condition, that is, in a compromised listening situation.

We thus interpret the positive correlations between gray-matter probability in the IPL and optimal cue utilization as evidence for its role in adaptive auditory categorization. More gray-matter probability in this substructure of the executive network is beneficial for categorization, because this area supports improved auditory selective attention, particularly relevant for compromised listening situations. Although the existing literature

suggests that the parietal cortex supports selective attention independent of stimulus modality, our data do not allow for a conclusive answer of whether the identified network responds to cognitive flexibility in a yet more general sense. Regarding the data of Henry et al. (in press), the implication of the right IPL might depend on whether attention is directed to or away from stimulus duration. It is not clear whether a switch in cue utilization would be similarly supported by the IPL if the acoustic cues were not related to durational aspects. Clearly, this issue must be approached by future research.

The tendency to utilize temporal cues after spectral degradation was also predicted by gray-matter probability in the supra-marginal gyrus (SMG, part of BA 40). This area is often found to be involved in auditory processing (e.g. Gaab, Gaser, & Schlaug, 2006; Jacquemot, Pallier, LeBihan, Dehaene, & Dupoux, 2003; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008; Obleser, Wöstmann, Hellbernd, Wilsch, & Maess, 2012). Arguably, the SMG supports executive control and attentional processing rather than domain-specific, auditory processing (Cabeza, Ciaramelli, & Moscovitch, 2012). More precisely, the SMG has been identified to be involved in orienting attention to stimulus aspects that are crucial for task performance (for durational aspects, cf. Livesey, Wall, & Smith, 2007; Wiener, Hamilton, Turkeltaub, Matell, & Coslett, 2010; Wiener, Turkeltaub, & Coslett, 2010). This provides further evidence for our interpretation that increased gray-matter probability in inferior parietal lobule is beneficial for adaptively re-directing attention to the most informative stimulus cue (i.e., duration under spectral degradation).

More generally, our data also speak to the ongoing discussion about potential neural predispositions that may be beneficial for second-language learning. As alluded to in the introduction, prior work on selective attention to acoustic cues has had a strong phonetic focus (Francis et al., 2008) and has particularly investigated how selective attention may help to acquire novel phonetic categories (Francis et al., 2000; Francis & Nusbaum, 2002). The present acoustic stimuli had spectral and temporal characteristics that are relevant for distinguishing, for example, vowel duration and vowel quality (Stevens, 1998). In this respect, our study suggests new avenues for studying the neural predispositions that support the redirection of attention from spectral cues to temporal cues, as is, for instance, required for a German learner of Finnish vowel duration (Ylinen et al., 2010).

Note further that our data do not speak to any causal direction in the correlations between gray-matter probability and listening behavior: from a developmental point of view, it is a question of future longitudinal research to specify whether increased gray-matter probability in parietal cortex causes more successful coping with demanding listening situations, or whether long-term adaptive behavior had driven this increase in gray-matter probability.

Finally, our VBM analyses revealed that gray-matter probability in precentral gyrus was predictive of optimal cue switch. We argue that the morphology in this cluster, extending into parietal cortex and comprising primary motor and pre-motor cortex, predicts optimal categorization on the basis of its functional role in sensorimotor integration (e.g., Hickok, Houde, & Rong, 2011 for sensorimotor integration in speech). For optimal behavior in our categorization task, it was also necessary to correctly map each sensory stimulus to a specific motor response (cf. Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010), an issue that we will discuss in more detail in the next section.

4.2. Auditory thalamic morphology predicts overall categorization performance

Notably, overall performance (as measured by d') co-varied with brain morphology in auditory thalamus as well as in a

pre- and postcentral cortical cluster, comprising parts of (pre-)motor and somatosensory cortices. The positive correlation of d' with gray-matter probability in auditory thalamus (medial geniculate body; MGB) suggests that auditory categorization accuracy can be predicted by morphology in relatively low-level areas within the ascending auditory pathway. The importance of auditory thalamus for auditory processing has been highlighted by recent research that also showed crucial functional similarities between MGB and auditory cortex (Tang et al., 2012; Yvert et al., 2002). For instance, the MGB has been found to show stimulus-specific adaptation (Antunes, Nelken, Covey, & Malmierca, 2010), frequency tuning (Edeline & Weinberger, 1991), and sensitivity to fast-changing spectral information, particularly relevant for speech (von Kriegstein et al., 2008). In a similar vein, Díaz, Hintz, Kiebel, and Kriegstein (2012) provided evidence that dysfunction of the MGB is accompanied by difficulties in processing speech sounds. Altogether, the MGB is more than a simple gateway to the cortex; rather, it dynamically shapes incoming information for the enhancement of representations and perception of acoustic features in auditory cortex and higher areas (Frith & Friston, 1996; for review, see Bartlett, 2013). It may therefore also be considered as hub for cognitive processes for which thalamic-cortical and cortico-thalamic information exchange is essential.

Following from the positive correlation between gray-matter probability in auditory thalamus and performance in our auditory-categorization task, we hypothesize that increased thalamic gray matter is beneficial for various aspects of auditory categorization that are relevant for the task in our experiment: frequency selectivity (initial cue in the nondegraded condition; Lennartz & Weinberger, 1992), category distinction (attribution of category labels in our experiment; von Kriegstein et al., 2008) and sensitivity to attentional modulation (Frith & Friston, 1996).

Recent research has provided evidence for high connectivity of the MGB to motor and sensory areas (Behrens et al., 2003; Johansen-Berg et al., 2005). These connectivity findings are in line with the presently observed positive correlation of MGB gray-matter probability with overall categorization performance. Note that our experiment comprised a strong sensorimotor component: participants had to learn sound-motor associations, in that they had to press one button for a tone of category A and another button for a tone of category B (cf. Maddox et al., 2010). Thus, in order to show good performance in this task, it was crucial that sensorimotor integration was successful. We therefore hypothesize that increased gray-matter probability in somatosensory (BA 3), motor (BA 4) and pre-motor (BA 6) cortex is beneficial for sensorimotor integration, and thus, for a successful performance in the auditory-categorization task. This interpretation is corroborated by studies demonstrating the importance of pre-motor areas for sensorimotor integration (Bangert et al., 2006; Chen, Penhune, & Zatorre, 2009; Rauschecker, 2011; Schubotz & von Cramon, 2002). We are aware, however, that future studies are necessary in order to exclude possible task-dependent involvement of motor areas. Possibly, experiments without task-related motor involvement might not show correlations in these areas.

Somewhat surprisingly, we found that gray-matter probability in primary and secondary auditory cortices, and generally, in (superior) temporal regions, was not predictive of auditory categorization and cue utilization in this experiment. Likely, our task had a strong focus on cue utilization and the enhancement of cue-relevant acoustic properties by subcortical structures (i.e., the MGB, cf. Frith & Friston, 1996). Therefore, the most predictive structural dispositions for successful categorization were found in auditory thalamus and in the parietal attention network, rather than in temporal auditory areas. The lack of an effect in auditory cortex might reflect the possibility that thalamic structural dispositions are more important for short-term auditory learning

(on the order of several minutes), while auditory cortical structural dispositions would be more relevant for long-term auditory learning (on the order of several weeks or months). Indeed, existing studies point in that direction (Edeline & Weinberger, 1991; Floyer-Lea & Matthews, 2005; Skoe, Krizman, Spitzer, & Kraus, 2013). In particular, Skoe et al. (2013) provided evidence that structures at low levels of the auditory pathway indeed support very fast auditory learning (within several minutes), suggesting that in general, hierarchically low-level areas in the auditory pathway are particularly suited for rapid learning (cf. Erb et al., 2012). Future research is necessary in order to examine whether dynamic measure of brain activity (e.g., fMRI) would show a similar focus on subcortical and parietal networks, or would alternatively demonstrate the involvement of primary and secondary auditory cortices.

4.3. Conclusions

The present findings provide insights into the brain structural prerequisites for adaptive listening behavior. Gray-matter probability in subcortical thalamic auditory networks as well as in cortical motor- and sensory-networks was shown to be predictive of auditory categorization performance. A switch from initial spectral cue utilization to temporal cue utilization under spectral degradation reflected adaptive categorization behavior. Importantly, we found that parts of a domain-general, parietal attention network in inferior parietal lobule were predictive of this switch. In sum, the results provide an important link in understanding how brain morphology in domain-specific and domain-general areas co-varies with human processing of complex sensory input and perceptual adaptation.

Acknowledgments

Research was supported by the Max Planck Society (Max Planck Research Group fund to J.O.). We thank two anonymous reviewers for their thoughtful contributions.

Appendix A. Supplementary Materials

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.neuropsychologia.2013.09.012>.

References

Alain, C., & Arnott, S. R. (2000). Selectively attending to auditory objects. *Frontiers in Bioscience: A Journal and Virtual Library*, 5, D202–212.

Anderson, J. S., Ferguson, M. A., Lopez-Larson, M., & Yurgelun-Todd, D. (2010). Topographic maps of multisensory attention. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 20110–20114.

Antunes, F. M., Nelken, I., Covey, E., & Malmierca, M. S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS One*, 5.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage*, 38, 95–113.

Ashburner, J. (2009). Computational anatomy with the SPM software. *Magnetic Resonance Imaging*, 27, 1163–1174.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry: The methods. *Neuroimage*, 11, 805–821.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., et al. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *Neuroimage*, 30, 917–926.

Bartlett, E. L. (2013). The organization and physiology of the auditory thalamus and its role in processing acoustic features important for speech perception. *Brain and Language*, 126, 29–48.

Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., et al. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6, 750–757.

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436, 1161–1165.

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex*, 19, 1583–1596.

Bermudez, P., & Zatorre, R. J. (2005). Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, 1060, 395–399.

Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, 7, 295–301.

Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*, 17, 1353–1366.

Brunetti, M., Della Penna, S., Ferretti, A., Del Gratta, C., Cianflone, F., Belardinelli, P., et al. (2008). A frontoparietal network for spatial attention reorienting in the auditory domain: A human fMRI/MEG study of functional and temporal dynamics. *Cerebral Cortex*, 18, 1139–1147.

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, 12, 679–690.

Cabeza, R., Ciaramelli, E., & Moscovitch, M. (2012). Cognitive contributions of the ventral parietal cortex: An integrative theoretical account. *Trends in Cognitive Sciences*, 16, 338–352.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences*, 1169, 15–34.

Cohen, Y. E. (2009). Multimodal activity in the parietal cortex. *Hearing Research*, 258, 100–105.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, 3, 292–297.

Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience*, 17, 641–651.

Díaz, B., Hintz, F., Kiebel, S. J., & Kriegstein, K. v. (2012). Dysfunction of the auditory thalamus in developmental dyslexia. *Proceedings of the National Academy of Sciences*, 109, 13841–13846.

Druilman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *Journal of the Acoustical Society of America*, 95, 1053–1064.

Edeline, J. M., & Weinberger, N. M. (1991). Subcortical adaptive filtering in the auditory system: Associative receptive field plasticity in the dorsal medial geniculate body. *Behavioral Neuroscience*, 105, 154–175.

Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2012). Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia*, 50, 2154–2164.

Floyer-Lea, A., & Matthews, P. M. (2005). Distinguishable brain activation networks for short- and long-term motor skill learning. *Journal of Neurophysiology*, 94, 512–518.

Foster, N. E. V., & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *Neuroimage*, 53, 26–36.

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception and Psychophysics*, 62, 1668–1680.

Francis, A. L., Kaganovich, N., & Driscoll-Huber, C. (2008). Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *Journal of the Acoustical Society of America*, 124, 1234–1251.

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology Human Perception and Performance*, 28, 349–366.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91, 1357–1392.

Frith, C. D., & Friston, K. J. (1996). The role of the thalamus in “top down” modulation of attention to sound. *Neuroimage*, 4, 210–215.

Gaab, N., Gaser, C., & Schlaug, G. (2006). Improvement-related functional plasticity following pitch memory training. *Neuroimage*, 31, 255–263.

Gaser, C., & Schlaug, G. (2003). Gray matter differences between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, 999, 514–517.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, 16, 129–135.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 47, 103–138.

Golestani, N., Price, C. J., & Scott, S. K. (2011). Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *Journal of Neuroscience*, 31, 4213–4220.

Goudbeek, M., Swingle, D., & Smits, R. (2009). Supervised and unsupervised learning of multidimensional acoustic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1913–1933.

Griffiths, T. D. (2001). The neural processing of complex sounds. *Annals of the New York Academy of Sciences*, 930, 133–142.

Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience*, 4, 633–637.

Guenther, F. H., Nieto-Castanon, A., Ghosh, S. S., & Tourville, J. A. (2004). Representation of sound categories in auditory cortical maps. *Journal of Speech, Language, and Hearing Research*, 47, 46–57.

Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 12, 140–149.

- Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage*, 22, 676–687.
- Henry, M. J., Herrmann, B., & Obleser, J. Selective attention to temporal features on nested time scales. *Cerebral Cortex*, <http://dx.doi.org/10.1093/cercor/bht240>, in press.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, 69, 407–422.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, 8, 393–402.
- Hill, K. T., & Miller, L. M. (2010). Auditory attentional control and selection during cocktail party listening. *Cerebral Cortex*, 20, 583–590.
- Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119, 3059–3071.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*, Vol. 354. New York: Wiley.
- Husain, F. T., Fromm, S. J., Pursley, R. H., Hosey, L. A., Braun, A. R., & Horwitz, B. (2006). Neural bases of categorization of simple speech and nonspeech sounds. *Human Brain Mapping*, 27, 636–651.
- Hutchison, E. R., Blumstein, S. E., & Myers, E. B. (2008). An event-related fMRI investigation of voice-onset time discrimination. *NeuroImage*, 40, 342–352.
- Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 23, 9541–9546.
- Johansen-Berg, H., Behrens, T. E. J., Sillery, E., Ciccarelli, O., Thompson, A. J., Smith, S. M., et al. (2005). Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cerebral Cortex*, 15, 31–39.
- Kiefer, M., Sim, E. -J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience*, 28, 12224–12230.
- Lennartz, R. C., & Weinberger, N. M. (1992). Frequency selectivity is related to temporal processing in parallel thalamocortical auditory pathways. *Brain Research*, 583, 81–92.
- Leung, A. W. S., & Alain, C. (2011). Working memory load modulates the auditory “what” and “where” neural networks. *NeuroImage*, 55, 1260–1269.
- Liebenthal, E., Desai, R., Ellingson, M. M., Ramachandran, B., Desai, A., & Binder, J. R. (2010). Specialization along the left superior temporal sulcus for auditory categorization. *Cerebral Cortex*, 20, 2958–2970.
- Lipski, S. C., Escudero, P., & Benders, T. (2012). Language experience modulates weighting of acoustic cues for vowel perception: An event-related potential study. *Psychophysiology*, 49, 638–650.
- Livesey, A. C., Wall, M. B., & Smith, A. T. (2007). Time perception: Manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia*, 45, 321–331.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Erlbaum.
- Maddox, W. T., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby, F. G. (2010). Category label and response location shifts in category learning. *Psychological Research—Psychologische Forschung*, 74, 219–236.
- Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., & Chenevert, T. L. (2000). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *Journal of Cognitive Neuroscience*, 12(Suppl. 2), 130–144.
- Myers, E. B. (2007). Dissociable effects of phonetic competition and category typicality in a phonetic categorization task: An fMRI investigation. *Neuropsychologia*, 45, 1463–1473.
- Myers, E. B., Blumstein, S. E., Walsh, E., & Eliassen, J. (2009). Inferior frontal regions underlie the perception of phonetic category invariance. *Psychological Science*, 20, 895–903.
- Myers, E. B., & Swan, K. (2012). Effects of category learning on neural sensitivity to non-native phonetic categories. *Journal of Cognitive Neuroscience*, 24, 1695–1708.
- Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, 32, 12376–12383.
- Peelle, J. E., Cusack, R., & Henson, R. N. A. (2012). Adjusting for global effects in voxel-based morphometry: Gray matter decline in normal aging. *Neuroimage*, 60, 1503–1516.
- Posner, M. I., & Dehaene, S. (1994). Attentional networks. *Trends in Neurosciences*, 17, 75–79.
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271, 16–25.
- Raz, A., & Buhle, J. (2006). Typologies of attentional networks. *Nature Reviews. Neuroscience*, 7, 367–379.
- Rinne, T., Stecker, G. C., Kang, X., Yund, E. W., Herron, T. J., & Woods, D. L. (2007). Attention modulates sound processing in human auditory cortex but not the inferior colliculus. *Neuroreport*, 18, 1311–1314.
- Rosen, S., Faulkner, A., & Wilkinson, L. (1999). Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *Journal of the Acoustical Society of America*, 106, 3629–3636.
- Salmi, J., Rinne, T., Koistinen, S., Salonen, O., & Alho, K. (2009). Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Research*, 1286, 155–164.
- Scharinger, M., Henry, M. J., & Obleser, J. (2013). Prior experience with negative spectral correlations promotes information integration during auditory category learning. *Memory and Cognition*, 41, 752–768.
- Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, 5, 688–694.
- Schubotz, R. I., & von Cramon, D. Y. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *NeuroImage*, 15, 787–796.
- Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*, 1252, 229–236.
- Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, 26, 100–107.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303–304.
- Sharda, M., & Singh, N. C. (2012). Auditory perception of natural sound categories – An fMRI study. *Neuroscience*, 214, 49–58.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Fulbright, R. K., Skudlarski, P., Mencl, W. E., et al. (2001). The functional neural architecture of components of attention in language-processing tasks. *Neuroimage*, 13, 601–612.
- Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine*, 31, 241–253.
- Shergill, S. S., Bullmore, E., Simmons, A., Murray, R., & McGuire, P. (2000). Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *American Journal of Psychiatry*, 157, 1691–1693.
- Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, 243, 104–114.
- Smits, R., Sereno, J., & Jongman, A. (2006). Categorization of sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 733–754.
- Stevens, K. N. (1998). *Acoustic phonetics*, Vol. 30. Cambridge, MA; London, England: The MIT Press.
- Tang, J., Yang, W., & Suga, N. (2012). Modulation of thalamic auditory neurons by the primary auditory cortex. *Journal of Neurophysiology*, 108, 935–942.
- Teki, S., Chait, M., Kumar, S., von Kriegstein, K., & Griffiths, T. D. (2011). Brain bases for auditory stimulus-driven figure-ground segregation. *Journal of Neuroscience*, 31, 164–171.
- Turkeltaub, P. E., & Coslett, H. B. (2010). Localization of sublexical speech perception components. *Brain and Language*, 114, 1–15.
- von Kriegstein, K., Patterson, R. D., & Griffiths, T. D. (2008). Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. *Current Biology*, 18, 1855–1859.
- Wang, X., & Humes, L. E. (2008). Classification and cue weighting of multidimensional stimuli with speech-like cues for young normal hearing and elderly hearing-impaired listeners. *Ear and Hearing*, 29, 725–745.
- Wiener, M., Hamilton, R., Turkeltaub, P., Matell, M. S., & Coslett, H. B. (2010). Fast forward: Supramarginal gyrus stimulation alters time measurement. *Journal of Cognitive Neuroscience*, 22, 23–31.
- Wiener, M., Turkeltaub, P. E., & Coslett, H. B. (2010). Implicit timing activates the left inferior parietal cortex. *Neuropsychologia*, 48, 3967–3971.
- Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 676–681.
- Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., et al. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *Journal of Cognitive Neuroscience*, 22, 1319–1332.
- Yvert, B., Fischer, C., Guénou, M., Krolak-Salmon, P., Isnard, J., & Pernier, J. (2002). Simultaneous intracerebral EEG recordings of early auditory thalamic and cortical activity in human. *European Journal of Neuroscience*, 16, 1146–1150.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6, 37–46.