PERGAMON

Neuropsychologia 40 (2002) 1956–1964

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

The neuroanatomical substrate of sound duration discrimination

Pascal Belin [a,b,*], Stephen McAdams [c], Lionel Thivard [a], Bennett Smith [c], Sophie Savel [d], Monica Zilbovicius [a,e], Séverine Samson [f], Yves Samson [a,g]

[a] *Groupe de Neurologie, CEA-SHFJ, DRM, 4 place du Général Leclerc, F-91406 Orsay Cedex, France*
[b] *Groupe de Recherche en Neuropsychologie Expérimentale, Département de Psychologie, Université de Montréal, Montreal, Que., Canada H3C 3J7*
[c] *Institut de Recherche et Coordination Acoustique/Musique (IRCAM-Centre National de la Recherche Scientifique), F-75004 Paris Cedex, France*
[d] *Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, F-75006 Paris Cedex, France*
[e] *ERM 0205–CEA-SHFJ, DSV, DRM, Orsay Cedex, France*
[f] *Université Charles de Gaulle Lille III, BP 149, F-59653 Villeneuve d'Asq Cedex, France*
[g] *Service des Urgences Cérébro-Vasculaires, AP-HP Hôspital de La Salpêtrière, F-75651 Paris Cedex 13, France*

Received 23 March 2001; received in revised form 15 April 2002; accepted 24 April 2002

Abstract

We investigated the neuroanatomical substrate of sound duration discrimination, using the same experimental design as in a previous study on sound intensity discrimination [J. Neurosci. 18 (16) (1998) 6388]. Seven normal subjects were trained to detect deviant sounds presented with a slightly longer duration than a 300 ms long standard harmonic sound, using a Go/No Go paradigm. Individual psychometric curves were assessed using a three-step psychoacoustic procedure. Subjects were then scanned while passively listening to the standard sound, and while discriminating changes in sound duration at four different performance levels ($d' = 1.5$, 2.5, 3.5 and 4.5). Analysis of regional cerebral blood flow (rCBF) data outlined activation, during the discrimination conditions, of a right hemispheric fronto-parietal network very similar to the one previously observed for intensity discrimination, as well as additional activation in the right prefrontal cortex (Brodmann Area (BA) 10), bilateral basal ganglia and cerebellar hemispheres. These findings suggest that discrimination of sound duration, as for discrimination of sound intensity, involves two cerebral networks: a supramodal right fronto-parietal cortical network responsible for allocation of sensory attentional resources, and a network of regions such as the basal ganglia, cerebellum, and right prefrontal cortex, more specifically involved in the temporal aspects of the discrimination task.

Keywords: Audition; Temporal processing; Perception; Attention; Psychoacoustics; Neuroimaging

1. Introduction

In a previous report, we described the results from a study that combined psychoacoustics and functional neuroimaging to investigate the neuroanatomical substrate of sound intensity discrimination in humans [2]. In that study, subjects were trained to detect deviant sounds presented with a slightly higher intensity than a standard harmonic sound, using a Go/No Go paradigm, for which individual psychometric curves were assessed. They were then scanned while passively listening to the standard sounds and while discriminating changes in sound intensity in a Go/No Go paradigm, at four different performance levels ($d' = 1.5$, 2.5, 3.5 and 4.5). Analysis of regional cerebral blood flow (rCBF) data showed activation of a right-hemispheric fronto-parietal network, presumably involved in allocation of supramodal sensory attentional resources, and of a region of secondary auditory cortex presumably involved in sensory computation of sound intensity differences [2]. In the present study, the same design was used to focus on discrimination of sound duration.

Duration is an important attribute of sensory stimuli in general, and of auditory stimuli in particular. Aside from its influence on perceived sound loudness for values under 200 ms [36], sound duration often plays a significant ecological role. For example, mating signals of different frog species can be spectrally very similar, and distinguishable only by temporal criteria such as pulse repetition rate, or in durational terms inter-pulse period [17]. In human auditory communication, duration of specific parts of our complex auditory signals usually carry meaningful information. Temporal information is especially important in speech, where it provides phonetic and prosodic cues at the levels of the envelope and of the periodicity, as well as in its fine structure [41]. Certain phonetic boundaries are defined only by the

* Corresponding author. Tel.: +1-514-343-2330; fax: +1-514-343-5787.
E-mail address: pascal.belin@umontreal.ca (P. Belin).

0028-3932/02/$ – see front matter © 2002 Published by Elsevier Science Ltd.
PII: S0028-3932(02)00062-3

duration of specific portions of the vocalization: for example, the main difference between /ta/ and /da/ lies in the duration, in the tens of milliseconds range, of the unvoiced part (voice onset time) that follows initial burst release before vowel onset [6]. On a longer time-scale, the suprasegmental, prosodic organization of speech utterances also contains crucial temporal information: subtle meanings and emotions are conveyed by minute differences in duration of pauses or syllables in conversations or calls [32], or in duration of notes and silences in musical interpretation. Discrimination of sound duration thus plays an important ecological role.

At the neurophysiological level, experimental data in animals and humans suggest that temporal cognition involves several central nervous structures. In the cat, discrimination of tone duration cannot be relearned after bilateral ablation of auditory cortical areas and degeneration of corresponding areas of the medial geniculate bodies [42]. In rats conditioned to expect reinforcement at a given delay after the conditioning stimulus, ablation of frontal cortex appears to change the reference memory for the expected time of reinforcement, so that they expect reinforcement later than it actually occurs [33]. Neurons in the frog's auditory thalamus show evoked responses critically dependent on duration of individual tones [15]. Similarly, recent electrophysiological evidence in the little brown bat indicates that most auditory cortical neurons respond preferentially to a narrow range of stimulus durations (some being tuned to a 'best duration'), with those neurons showing long-pass and short-pass duration response functions being narrowly distributed within two narrow rostrocaudal slabs [10].

In humans, the different aspects of temporal cognition seem to be subserved by a complex network of interconnected cortical and subcortical cerebral structures, including the cerebellum, basal ganglia, and several regions of the cortex, particularly in the parietal and frontal lobes of the right hemisphere [21,18,27,12,16,26,40], for a review see [28]. At the level of the auditory cortex, several studies also suggest that discrete neural populations of the supratemporal plane might be responsive to changes in sound duration. Patients with resection within the temporal lobes for relief of intractable epilepsy were found to be impaired in the subtest of the Seashore Measures of Musical Talents that focuses on perception of sound duration, and impairments were significantly stronger for patients with right temporal lobe resections [31]. More recently, studies using electro- or magneto-encephalography have found that changes in sound duration elicit a mismatch negativity (MMN) generated by neuronal populations of the superior temporal plane [22,11], with stronger response on the right side [11]. Jaramillo et al. [20] also found that the MMN was elicited by both decrements and increments in duration, and that, for most durations tested, it increased in amplitude as a function of the amount of deviation from the standard duration [20]. Recently, Rao et al. [40] used functional magnetic resonance imaging (fMRI) to measure brain activity during an auditory time perception task where subjects had to judge the duration of a time interval defined by short tones, relative to a standard 1200 ms interval. Two control conditions with a similar design, including a pitch judgment task, revealed activity specifically related to time perception processes. Moreover, the temporal resolution of fMRI allowed the authors to distinguish processes associated with encoding time intervals from those related to comparing intervals [40].

The design of the present study was the same as for the previous study on intensity discrimination, except that the deviant sounds were slightly *longer* than the 300 ms long standard sound instead of being slightly *louder*. Each subject was scanned using positron emission tomography (PET) while passively listening to the standard sounds presented at regular 1 s intervals, and while covertly detecting longer deviants (25% probability of occurrence) among the standard sounds. This Go/No Go task was performed at four levels of performance, as defined by an unbiased decision parameter derived from signal detection theory ($d' = 1.5$, 2.5, 3.5 and 4.5). The goal of this study was to investigate the neural substrate of auditory temporal discrimination in humans, and compare the results to those of similar studies in other sensory modalities. We were also interested in how attending to two different attributes of auditory stimuli (intensity and duration) would affect the pattern of neuronal activity, measured in two groups of subjects under the same experimental conditions.

2. Materials and methods

2.1. Subjects

Seven healthy male volunteers (aged 22–33 years) gave written informed consent. They had no history of neurological or psychiatric disorders and had normal hearing. The study was approved by the ethics committee of the La Salpêtrière Hospital.

2.2. Auditory stimuli

Auditory stimuli were synthesized at a sampling rate of 44.1 kHz using the IRCAM Musical Workstation (ISPW digital signal processing card and a NeXT computer). The stimuli were harmonic complexes with 20 harmonics and a fundamental frequency of 200 Hz. The relative amplitudes of the harmonics were determined by a $1/n$ spectral envelope where n is the harmonic rank (−3 dB per octave slope in the power spectrum). Each reference stimulus had a duration of 300 ms including 80 ms linear rise and decay ramps in the amplitude envelope. In each deviant stimulus, the rise and decay times remained constant and the steady-state portion was increased in duration. Stimuli were presented binaurally over earphones at a level of 75 dB SPL as measured with a Bruel and Kjaer 2209 sound level meter (a weighting, fast response).

2.3. Psychoacoustic measures

The main task used in the imaging studies was Go/No Go: in a series of events of which the majority (75%) have the reference duration and the minority (25%) have a duration which is greater by some chosen amount; the subject must decide whether each one is the reference value (in which case no action is taken, No Go) or a deviant value, which is always superior to the reference value (in which case, the change is noted mentally in the imaging task and a button is pushed in the psychoacoustic task, Go). Ideally, it should be possible to establish performance levels for this kind of task with varying duration differences. However, this task has not been studied much in human psychoacoustics, and the data cannot therefore be compared with the literature. The psychoacoustic measures were thus made using a classic same/different task with an adaptive method (Phase 1), with a method of constant stimuli (Phase 2), as well as the Go/No Go task (Phase 3). In the same/different tasks, each trial was composed of two sounds presented sequentially. Four combinations were possible: two reference (R) stimuli (same), two test (T) stimuli (same), and one of each in the two orders (different).

2.3.1. Phase 1: preliminary measure of high and low thresholds by an adaptive method

To reduce experimentation time, a first rough measure of each subject's sensitivity to duration change was made using a *N*-down, 1-up adaptive procedure [23] which converges on a performance level that depends on *N*. *N* consecutive correct responses result in a decrease in duration difference and 1 incorrect response results in an increase. In our case, *N* was 3 (79.4%, low threshold, TL) and 8 (91.7%, high threshold, TH). There were four trial types: two reference stimuli (RR, same), two test stimuli (TT, same) and one of each in the two orders (RT and TR, different). On each trial, one of the four trial types was chosen randomly. The durations of T stimuli at which the adaptive trajectory changes direction were recorded: the last 8 of 12 were averaged to estimate TL, and the last four of six were averaged for TH. Six estimates were obtained for each threshold. From the mean duration differences obtained at each threshold for each subject, the values used in Phase 2 were determined. If half the duration difference between TL and TH is denoted *d*, the tested durations included TL $- d$, TL, TL $+ d$, TH, TH $+ d$.

2.3.2. Phase 2: psychometric functions determined with the method of constant stimuli

For each of the five duration differences obtained in Phase 1, a block of 200 trials was constructed. There were 50 repetitions of each trial type (RR, TT, RT or TR) in the block, presented in random order. After hearing the two sounds the subject indicated whether the stimuli were the same or different. The hit rate was computed on "different" trials from the percentage of correct responses. The false alarm rate was computed on "same" trials from the percentage of incorrect responses. According to signal detection theory [14], the discrimination rate expressed as percentage of correct detection of duration change is influenced both by the subject's perceptual sensitivity and by his or her judgment strategy (or response bias). Since the interest of this study was to determine neural correlates of sensitivity to change in duration, the sensitivity (d') was estimated from hit and false alarm rates [25]. This d' value is considered to estimate true sensitivity to duration difference with biases due to response strategy having been factored out. This procedure was repeated for each duration difference in a random order for each subject. From the d' values estimated for each duration difference a psychometric function was determined from a linear regression of those d' values onto duration differences. The highest d' value was at times removed from the fitting procedure if the curve clearly asymptoted at maximum performance level.

2.3.3. Phase 3: psychometric functions determined with the Go/No Go method

On the basis of the previously determined psychometric function, five new duration differences were chosen for the Go/No Go procedure corresponding to performances in Phase 2 equivalent to d' values of 1, 2, 3, 4 and 5. Each difference was presented in a separate block of trials. A block lasted about 2 min as in the imaging experiment. During this time, 200 events were presented, of which 75% had the reference duration and 25% the (higher) test duration. The subject listened continuously and pressed a button as soon as a test event occurred. The events were presented at a rate of one per second. If the subject pressed the button during the 1 s temporal window corresponding to the test signal, it was scored as a hit. If the button press occurred outside of this window, it was scored as a false alarm. The d' values were estimated from hits and false alarm rates [25] as in Phase 2. The procedure was repeated for each of the five duration differences in a different random order for each subject. From the d' values for each of the five duration differences a linear psychometric curve was determined as in Phase 2 and duration differences corresponding to d' values of 1.5, 2.5, 3.5 and 4.5 were chosen for the imaging studies for each subject.

2.4. Functional neuroimaging

2.4.1. Scanning

Relative rCBF was determined from the distribution of radioactivity after bolus intravenous injections of $H_2{}^{15}O$ [7], measured with a ECAT-HR + PET camera (Siemens AG, Erlangen, Germany). Subjects received 12 $H_2{}^{15}O$ injections (9 mCi per injection) corresponding to 12 rCBF measurements, performed at 10 min intervals. Attenuation-corrected data were reconstructed into 63 2.25 mm thick axial slices, with a resulting resolution of 4.5 mm full-width-at-half-maximum after reconstruction [3].

2.4.2. Tasks

Four scans were acquired during a baseline condition, and eight scans during detection of changes in duration performed at four different levels of discriminability (two scans per level), in a counterbalanced order. The baseline condition consisted of passively listening to standard sounds, presented binaurally through Sony MDR-V600 headphones at a 75 dB sound pressure level, with a 1 s inter-onset interval. Subjects were informed that all sounds were identical and were instructed to listen carefully to the sounds. During the duration discrimination conditions, subjects were instructed to mentally detect sounds of longer duration (deviants), that were intermingled with the standard sounds of the baseline condition with a 25% probability of occurrence. The deviant sounds were identical to the standard sounds in all respects but duration, which was slightly greater in the deviants. During a given duration discrimination condition, all deviants were identical, with a duration corresponding to a given performance level ($d' = 1.5$, 2.5, 3.5 or 4.5) for the scanned subject. In order to avoid possible contamination of the activation pattern by motor-specific activity, no overt response was required from the subjects. However, subjective duration change discriminability, as indicated by subjects after each scan, corresponded qualitatively to the individual's objective d' value, thus suggesting that they were performing the discrimination task as during the psychophysical sessions.

2.4.3. Analysis

Statistical parametric mapping (SPM96) software was used for image realignment, transformation into standard stereotaxic anatomical space [43], smoothing, and statistical analysis at each intra-cerebral volume element (2 mm × 2 mm × 2 mm) or voxel [8,9]. State-dependent differences in global flow were covaried out using proportional scaling. Comparisons across conditions were made using the *t*-statistic subsequently transformed into the normally distributed *Z* statistic (SPM(*Z*)). A categorical approach was first used to determine cerebral regions in which rCBF changed significantly in the seven subjects between baseline and the pooled discrimination conditions. A parametric approach was then used to determine regions in which rCBF showed a significant linear covariation with level of performance (d') during the duration discrimination conditions.

3. Results

3.1. Behavioral measures

Performance in the duration discrimination task varied a great deal across subjects for all three phases. This variation resulted in widely differing physical durations for a given d' level (Fig. 1). Fig. 1 also shows that the slopes of the psychometric functions were very similar, and that in all cases d' increased linearly with the duration of the deviant sound. The mean deviant duration discriminated from the reference 300 ms duration varied from 335 ms for a d' of 1.5 to 394 ms for a d' of 4.5. These values are consistent with a 30 ms discrimination threshold at 300 ms found by Abel [1] in human listeners. As with the previous intensity discrimination study [2], inter-subject variability was greater for higher d' levels (Fig. 1). The regression coefficients for the psychometric functions from which were derived the stimulus values used for the imaging study varied from 0.80 to 0.98 (mean = 0.94). The slopes for the psychometric functions derived from the constant stimuli phase were shallower than those from the Go/No Go phase indicating that improvement in sensitivity was still taking place during the last phase. Nonetheless, the use of psychoacoustic methods such as these ensured that the PET scan measured activation for four fixed levels of sensitivity in auditory discrimination of duration that were similar across subjects, in spite of differences in terms of the actual physical differences used.

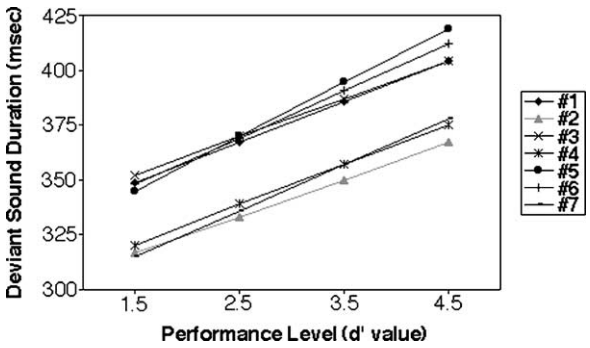

Fig. 1. Duration of the deviant sound discriminated from the 300 ms long standard, for each subject and each of four performance levels (d').

3.2. The rCBF variations with detection of duration changes

When compared with the baseline, the four detection conditions averaged together yielded significant rCBF increases ($P < 10^{-6}$ uncorrected, $P < 0.05$ corrected for multiple comparisons) in several regions of the deep gray nuclei, in the cerebral cortex of the right hemisphere, and in the cerebellum bilaterally (Table 1; Fig. 2). Highest activation was observed in an extensive part of the right prefrontal cortex, peaking in a circumscribed region of the inferior frontal operculum (Brodmann Area (BA) 45); it extended anteriorly to right orbital prefrontal cortex (BA 10) and medially to subcortical structures bilaterally, with an important maximum of activation in the right thalamus. Intense rCBF changes were also observed in the left cerebellum, peaking in the superior and inferior parts of the left cerebellar hemisphere, and these extended to the superior part of the right cerebellar hemisphere (Table 1). Additional rCBF increases were found in the right inferior

Table 1
Brain activity associated with sound duration discrimination compared with the baseline

	x	*y*	*z*	*Z*-value	Size
rCBF increases					
Right inferior frontal operculum (45)	38	18	8	6.25	11506
Right dorsomedial thalamus	6	−12	8	6.15	
Right orbital prefrontal (10)	24	52	−8	5.92	
Right putamen	24	6	4	5.57	
Right caudate nucleus	12	12	2	5.53	
Right inferior frontal gyrus (44)	32	8	26	5.17	
Right posterior thalamus	6	−24	14	5.00	
Right inferior frontal gyrus (44)	42	8	18	5.00	
Right middle frontal gyrus (10)	36	38	14	4.98	
Right middle frontal gyrus (46)	40	30	24	4.95	
Right middle frontal gyrus (46/10)	32	42	6	4.86	
Left lateral cerebellum (VIII A–L)	−28	−60	−48	6.17	5018
Left lateral cerebellum (VI L)	−30	−62	−28	6.13	
Left posterior cerebellum (VII A)	−18	−82	−34	5.39	
Left cerebellum, center	−14	−64	−38	5.37	
Left anterior cerebellum (VI L)	−20	−48	−32	4.95	
Right lateral cerebellum, lobule (VI L)	32	−58	−24	5.71	827
Right middle temporal gyrus (21)	66	−30	−2	4.93	265
Right inferior parietal lobule (40)	54	−50	50	4.78	227
rCBF decreases					
Left inferior parietal lobule (39)	−42	−70	24	5.81	1471
Left precuneus, superior (7)	−4	−56	56	5.56	1318
Left precuneus, inferior (7)	−4	−58	34	4.89	

Coordinates (in standard stereotaxic space [43]) of voxels corresponding to local maxima of *Z*-value, above $Z = 4.75$ ($P < 10^{-6}$) within each focus of activation; *x*: distance (mm) to right (+) or left (−) of the mid-sagittal line, *y*: distance anterior (+) or posterior (−) to the vertical plane through the anterior commissure, *z*: distance above (+) or below (−) the inter-commissural (AC-PC) line. Approximate Brodmann numbers (BA) associated with anatomical regions are given in parentheses. Cerebellar lobules, determined using the probabilistic atlas described in [5], are given in parentheses. Size refers to the number of voxels in a given cluster (voxel size: 2 mm × 2 mm × 2 mm), for SPM(*Z*) maps thresholded at $Z = 3.72$ ($P < 0.0001$ uncorrected), and then corrected for multiple non-independent comparisons at $P < 0.05$.

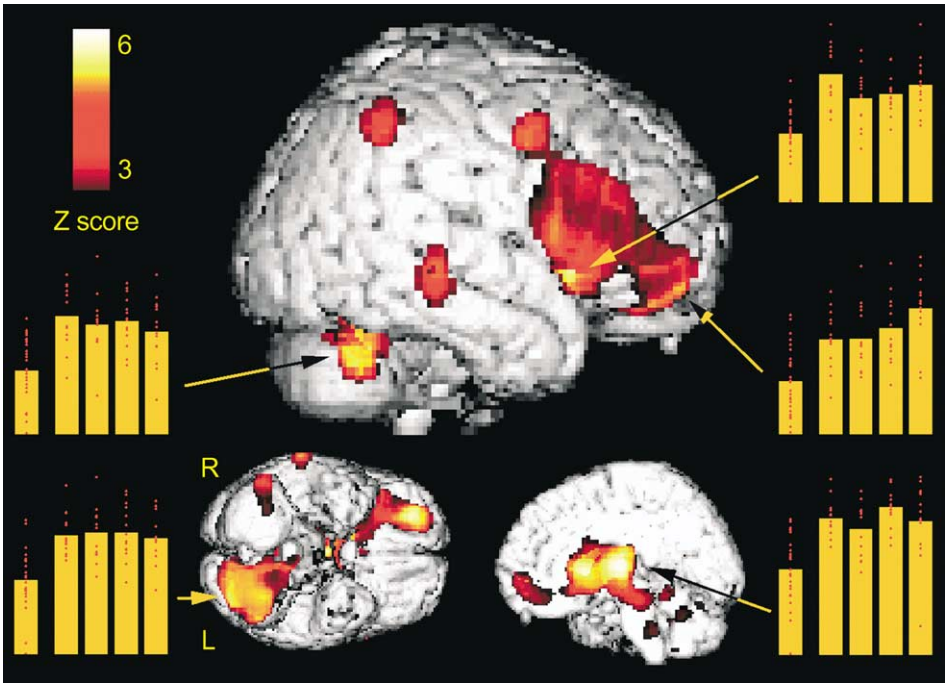

Fig. 2. Regions significantly more active during discrimination of sound duration (all performance levels pooled) than during passive listening to the standard sound are rendered on a cortical surface viewed from the right (center), from the bottom (lower left), and on the medial surface of a right hemisphere (bottom right). The bar diagrams represent, for each of these regions, the mean (yellow bar) and individual (red dots) rCBF values corresponding to the baseline (left bar) and each of the four duration discrimination conditions (right bars, $d' = 4.5$, 3.5, 2.5 and 1.5 from left to right), in arbitrary units.

Table 2
Regions of significant correlation between rCBF and performance level as measured by d' value

	x	y	z	Z-value	Size
Negative rCBF–d' correlation					
Left inferior parietal lobule (39/40)	−48	−66	40	3.50	55
Left temporal pole (38/20)	−48	10	−38	3.50	127
Left temporal pole (20)	−44	−2	−36	3.42	
Right middle temporal gyrus (21)	58	−8	−18	3.28	209

See the legend of Table 1. SPM(Z) maps were thresholded at $Z = 2.33$ ($P < 0.001$ uncorrected for multiple, non-independent correlations).

parietal lobule (BA 40) and the right middle temporal gyrus (BA 21). Discrimination conditions also yielded significant rCBF decreases when compared to the passive listening baseline, restricted to the posterior part of the left hemisphere, and located in inferior parietal lobule and precuneus (Table 1).

3.3. rCBF: performance correlation

A parametric approach was used to determine possible cerebral regions in which normalized activity was linearly related with equivalent performance level—as measured by d' value—during the detection of changes in sound duration. Three regions were found to show a significant ($P < 0.001$ uncorrected) negative correlation between rCBF and d' value; in these regions, rCBF increased with task difficulty (Table 2). Conversely, no regions showed the opposite pattern of a significant ($P < 0.001$ uncorrected) positive correlation between rCBF and d' value, a pattern similar to the one observed in the case of intensity discrimination. These correlations should be, however, interpreted with caution since none of them reached the $P < 0.05$ criterion when corrected for multiple, non-independent comparisons, and they are based on a small number of psychophysical performance levels.

4. Discussion

Part of the activation pattern observed during duration discrimination was very similar to the one observed in the previous study on sound intensity discrimination [2]. It consisted of a set of fronto-parietal zones in the right hemisphere and left cerebellum regions whose collective involvement in sensory attention has been repeatedly suggested by clinical, electrophysiological and neuroimaging evidence [29,34,45,13,35]. This ‘attentional network’ can be described as composed of three main components, in which the neuronal activity is not necessarily co-varying: an extensive right prefrontal component, including large regions of the frontal operculum (BA 44, 45, 47) and extending caudally and dorsally to premotor regions in the vicinity of the frontal-eye-field (FEF) (BA 6); a region of the right inferior parietal lobule (BA 40) at approximately the same horizontal level as the FEF; and a large region of the left, contralateral cerebellar hemisphere. Activation in the right inferior frontal operculum peaked in the present study approximately 12 mm anteriorly to the peak found in the previous study ($x = 42$, $y = 28$, $z = 4$); in the right parietal lobe the peak of the present study was located 10 mm more lateral than the one previously reported ($x = 44$, $y = -48$, $z = 50$), and in the left cerebellum the peaks in these two groups of subjects were as close as 8 mm (previous study: $x = -38$, $y = -62$, $z = -26$). This similarity suggests that this attentional network might not be differently activated by the physical attribute on which the attention focused (e.g. intensity or duration).

A similar conclusion was proposed by Maquet et al. [27], who used a design very similar to the present one in the visual modality. In that study, subjects were scanned while they performed a two-alternative forced choice task in which they had to decide if the duration or intensity of a visual stimulus (green LED) matched that of a standard presented earlier [27]. When the duration judgment conditions were compared to a control condition of passive stimulation with the standard visual stimulus, greater activity was observed in all of the above three components. Although this study differed methodologically in several ways from the present one, it yielded activation in locations relatively close to those reported here for the right frontal ($x = 34$, $y = 20$, $z = 4$, 6 mm difference), right parietal ($x = 44$, $y = -52$, $z = 40$, 14 mm difference) and left cerebellar ($x = -16$, $y = -86$, $z = -28$, 6 mm difference) peaks of activation. Interestingly, such activation was found in both the duration judgment and the intensity judgment tasks (no significant differences were found between the two conditions in that study), again suggesting that the right fronto-parietal/left cerebellum activation could be activated in attentional tasks irrespective of the sensory modality and sensory feature being attended to.

In contrast with the previous study on intensity discrimination [2], however, no relation could be observed between neuronal activity in the right fronto-parietal network and level of discrimination performance. As seen in Fig. 2, level of activity in the regions associated with duration discrimination seemed largely independent of the performance level (d'), which is confirmed by the fact that regions with a significant correlation ($P < 0.001$) between CBF and d' (Table 2) did not show significantly different activity during baseline and discrimination conditions (Table 1). This unexpected finding could reflect fundamental differences in processing sensory attributes such as time and intensity, or differences across groups of subjects, and further work will be necessary to clarify the issue.

An important feature of the present results is the considerable subcortical activation induced by the duration discrimination task: significant activation was observed in the thalamus and basal ganglia, peaking in left thalamus

and extending to left and right striata (Table 1, Fig. 2). This result is consistent with several studies that emphasize the importance of the basal ganglia in timing operations [18,12,40]. In particular, Rao et al. [40] revealed with fMRI the early activation of the basal ganglia during an interval comparison task, suggesting an important role of these structures in encoding time intervals [40].

A second important feature of the present pattern of activation concerns the cerebellum: whereas neuroimaging studies report in general unilateral activation of the cerebellum, mostly contralateral to the site of main cortical activation, cerebellar activation in the present study was observed on both the sides (Table 1). In addition to the lateral and superior part of the lobule VI of the left cerebellar hemisphere [5], a nearly symmetrical region of the right cerebellum was activated only for duration discrimination ($x = 32$, $y = -58$, $z = -24$). This result is consistent with current models of temporal processing viewing the cerebellum, especially its lateral parts, as playing a critical role in timing operations [19,18]. A PET study of sound duration comparison [21] also emphasized the role of the cerebellum in timing operations in humans by showing bilateral activation of the cerebellar hemispheres in a duration comparison task. The activation foci they reported (left: $x = -14$, $y = -36$, $z = -12$; right: $x = 16$, $y = -76$, $z = -12$), however, were situated in a different location, higher and more medial than in the present study. Interestingly, a recent study of motor timing [38] reported cerebellar activation during production of a simple timed motor response that peaked in a very close location ($x = -29.5$, $y = -60.7$, $z = -31.5$, <4 mm difference) to the one activated bilaterally in the present study. Yet, Rao et al. [40] suggested that cerebellar activation in time perception task could be related to processes other than explicit timing, probably associated with the motor response. Importantly, the relatively high level of cerebellar activation observed in the present study should be related to the rather short range of durations used here (around 300 ms). Data from time perception experiments using longer durations and higher temporal resolution suggest that cerebellar activation typically occurs only around stimulus onset [28], compared to a more sustained activation of the basal ganglia.

A third feature of the present results concerns activation in the right prefrontal cortex: strong rCBF increases were found in right orbital prefrontal cortex (BA 10) for duration discrimination, in the ventral and anterior part of the superior frontal sulcus. This result is consistent with animal work suggesting a particular role of the frontal cortex for temporal discrimination [33], as well as with a recent study in brain-damaged human patients demonstrating the importance of right prefrontal cortex for the perception of duration [16]. Recent neuroimaging evidence in normal human subjects also emphasizes the involvement of right prefrontal cortex in temporal tasks [37,39]. Such activation could be specifically related to the comparison of time intervals [40].

Last, very few of the regions found to be more active during attention to sound duration than during the baseline were located in auditory-related areas of the cortex. Here, only one temporal lobe region was found to be activated during the discrimination task relative to the passive baseline, located in right middle temporal gyrus (Table 1), in a rather distant location from the anatomical site of primary auditory cortex; this result is consistent with the fact that auditory input was very similar in all conditions, with only slight duration increases for 25% of the stimuli in the discrimination conditions. Yet, specific attention to an auditory feature such as sound duration might have been expected to induce greater activation in auditory cortex than during non-focused, passive listening. This absence of attention-induced activation in primary auditory fields has already been observed [4,44,46,47], and is consistent with a model of cortical architecture according to which modulations become increasingly less pronounced in the stream from higher-order association areas to unimodal, secondary and primary cortices [30].

5. Conclusion

These results suggest that discrimination of sound intensity is performed in the human brain by the combination of two cortical networks, or groups of areas working in concert: (i) a sensory attentional ‘network’ [29,30] with components in the right frontal and parietal lobes, probably highly aspecific and supramodal since its activation is often reported for sensory attentional tasks in other modalities and for other sensory features; (ii) a temporal processing ‘network’ [18,12] involving, in particular, activation to varying degrees of the basal ganglia, cerebellar hemispheres and right prefrontal cortex (BA 10), presumably responsible for processing sound duration differences. While the different components of the fronto-parietal network seem to be recruited whenever attention to a sensory attribute is required, the extensive subcortical and right frontal activations observed here were probably specifically induced by the temporal component of the task.

For further reading see [24].

Acknowledgements

This work was supported by the Groupement d’Intérêt Scientifique-Sciences de la Cognition, Société d’Etudes et de Soins pour les Enfants Polymalformés and Fondation France-Télécom. We thank the staff members of the Institut de Recherche et Coordination Acoustique/Musique

and the Orsay Brain Imaging Center for technical assistance.

References

[1] Abel S. Duration discrimination of noise and tone bursts. Journal of the Acoustical Society of America 1972;51:1219–23.

[2] Belin P, McAdams S, Smith B, Savel S, Thivard L, Samson S, Samson Y. The functional anatomy of sound intensity discrimination. Journal of Neuroscience 1998;18:6388–94.

[3] Bendriem B, Casey M, Dahlbom M, Trebossen R, Blohm K, Nutt R, Syrota A. Evaluation of the ECAT EXACT HR+: a new positron camera with 2D/3D acquisition capabilities and nearly isotropic spatial resolution. In: Proceedings from the Society of Nuclear Medicine Annual Meeting, CO, 1996

[4] Binder J, Frost J, Hammeke T, Rao S, Cox R. Function of the left planum temporale in auditory and linguistic processing. Brain 1996;119:1239–47.

[5] Doyon J, Lavoie K, Penhune V, Collins L, Schmahmann JD, Evans AC, et al. 3D probabilistic mapping of the human cerebellum from MRI scans in stereotaxic space. Neuroimage 1999;9:S226.

[6] Fant G. Acoustic theory of speech production. The Hague: Mouton, 1960

[7] Fox PT, Mintun MA, Raichle ME, Herscovitch P. A non-invasive approach to quantitative functional brain mapping with $H_2{}^{15}O$ and positron emission tomography. Journal of Cerebral Blood Flow and Metabolism 1984;4:329–33.

[8] Friston K, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ. Spatial registration and normalization of images. Human Brain Mapping 1995;2:165–89.

[9] Friston K, Holmes AP, Worsley KJ, Poline J-B, Frith CD, Frackowiak RSJ. Statistical parametric mapping in functional imaging: a general linear approach. Human Brain Mapping 1995;2:189–210.

[10] Galazyuk AV, Feng AS. Encoding of sound duration by neurons in the auditory cortex of the little brown bat, *Myotis lucifugus*. Journal of Comparative Physiology 1997;180:301–11.

[11] Giard MH, Lavikainen J, Reinikainen K, Perrin F, Bertrand O, Pernier J, et al. Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. Journal of Cognitive Neuroscience 1995;7:113–43.

[12] Gibbon J, Malapani C, Dale CL, Gallistel C. Toward a neurobiology of temporal cognition: advances and challenges. Current Opinion in Neurobiology 1997;7:170–84.

[13] Gitelman DR, Alpert NM, Kosslyn S, Daffner K, Scinto L, Thompson W, et al. Functional imaging of human right hemispheric activation for exploratory movements. Annals of Neurology 1996;39:174–9.

[14] Green DM, Swets JA. Signal detection theory and psychophysics. New York: Krieger, 1997

[15] Hall J, Feng AS. Neural analysis of temporally patterned sounds in the frog's thalamus: processing of pulse duration and pulse repetition rate. Neuroscience Letters 1986;63:215–20.

[16] Harrington DL, Haaland KY, Knight RT. Cortical networks underlying mechanisms of time perception. Journal of Neuroscience 1998;18:1085–95.

[17] Hauser M. The evolution of communication. Cambridge (MA): MIT Press, 1996

[18] Ivry RB. The representation of temporal information in perception and motor control. Current Opinion in Neurobiology 1996;6:851–7.

[19] Ivry RB, Keele SW, Diener HC. Dissociation of the lateral and medial cerebellum in movement timing and movement execution. Experimental Brain Research 1988;73:167–80.

[20] Jaramillo M, Paavilainen P, Näätanen R. Mismatch negativity and behavioral discrimination in humans as a function of the magnitude of change in sound duration. Neuroscience Letters 2000;290:101–4.

[21] Jueptner M, Rijntjes M, Weiller C, Faiss JH, Timmann D, Mueller SP, et al. Localization of a cerebellar timing process using PET. Neurology 1995;45:1540–5.

[22] Kaukoranta E, Sams M, Hari R, Hamalainen M, Näätanen R. Reactions of human auditory cortex to a change in tone duration. Hearing Research 1989;41:15–21.

[23] Levitt H. Transformed up–down methods in psychoacoustics. Journal of the Acoustical Society of America 1977;49:467–77.

[24] Liegeois-Chauvel C, de Graaf JB, Laguitton V, Chauvel P. Specialization of left auditory cortex for speech perception in man depends on temporal coding. Cerebral Cortex 1999;9:484–96.

[25] Macmillan N, Creelman D. Detection theory: a user's guide. Cambridge: Cambridge University Press, 1990

[26] Mangels JA, Ivry RB, Shimizu N. Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. Cognitive Brain Research 1998;7:15–39.

[27] Maquet P, Lejeune H, Pouthas V, Bonnet M, Casini L, Macar F, et al. Brain activation induced by estimation of duration: a PET study. Neuroimage 1996;3:119–26.

[28] Mattel MS, Meck WH. Neuropsychological mechanisms of interval timing behaviour. Bioessays 2000;22:94–103.

[29] Mesulam M-M. A cortical network for directed attention and unilateral neglect. Annals of Neurology 1981;10:309–25.

[30] Mesulam M-M. From sensation to cognition. Brain 1998;121:1013–52.

[31] Milner B. Laterality effects in audition. In: Inter-hemispheric relations and cerebral dominance. Baltimore (MD): John Hopkins University Press, 1962. p. 177–95.

[32] Monrad-Krohn GH. The third element of speech: prosody and its disorders. In: Halpern L, editor. Problems of dynamic neurology. Jerusalem: Hebrew University Press, 1963. p. 101–17.

[33] Olton DS. Frontal cortex, timing and memory. Neuropsychologia 1989;27:121–30.

[34] Pardo JV, Fox PT, Raichle ME. Localization of a human system for sustained attention by positron emission tomography. Nature 1991;349:61–4.

[35] Paus T, Zatorre RJ, Hofle N, Caramanos Z, Gotman J, Petrides M, et al. Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. Journal of Cognitive Neuroscience 1997;9:392–408.

[36] Pedersen CB, Salomon G. Temporal integration of acoustic energy. Acta Otolaryngology 1977;83:417–23.

[37] Pedersen CB, Mirz F, Ovesen T, Ishizu K, Johannsen P, Madsen S, et al. Cortical centers underlying auditory temporal processing in humans: a PET study. Audiology 2000;39:30–7.

[38] Penhune VB, Zatorre RJ, Evans AC. Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. Journal of Cognitive Neuroscience 1998;10:752–65.

[39] Pouthas V, Garnero L, Fernadez AM, Renault B. ERPs and PET analysis of time perception: spatial and temporal brain mapping during visual discrimination tasks. Human Brain Mapping 2000;10:49–60.

[40] Rao SM, Mayer AR, Harrington DL. The evolution of brain activation during temporal processing. Nature Neuroscience 2001;4:317–23.

[41] Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. Philosophical Transactions of the Royal Society of London Series B 1992;336:367–73.

[42] Scharlock DP, Neff WD, Strominger NL. Discrimination of tone duration after bilateral ablation of cortical auditory areas. Journal of Neurophysiology 1965;28:673–81.

[43] Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. New York: Thieme, 1988

[44] Tzourio N, Massioui FE, Crivello F, Joliot M, Renault B, Mazoyer B. Functional anatomy of human auditory attention studied with PET. Neuroimage 1997;5:63–77.
[45] Woods RP, Cherry SR, Mazziota JC. Rapid automated algorithm for aligning and reslicing PET images. Journal of Computer Assisted Tomography 1992;16:620–33.
[46] Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch processing in speech perception. Science 1992;256:846–9.
[47] Zatorre RJ, Meyer E, Gjedde A, Evans AC. PET studies of phonetic processing of speech: review, replication and re-analysis. Cerebral Cortex 1996;6:21–30.