NeuroReport 10, 3825–3830 (1999)

A common neural substrate for the analysis of pitch and duration pattern in segmented sound?

Timothy D. Griffiths,[1,2,CA] Ingrid Johnsrude,[1] Jennifer L. Dean[2] and Gary G. R. Green[2]

[1]Wellcome Department of Cognitive Neurology, Institute of Neurology, 12 Queen Square, London, WC1N 3BG; [2]Department of Physiological Sciences, Newcastle University Medical School, Framlington Place, Newcastle-upon-Tyne, NE2 4HH, UK

[CA,2]Corresponding Author and Address

THE analysis of patterns of pitch and duration over time in natural segmented sounds is fundamentally relevant to the analysis of speech, environmental sounds and music. The neural basis for differences between the processing of pitch and duration sequences is not established. We carried out a PET activation study on nine right-handed musically naive subjects, in order to examine the basis for early pitch- and duration-sequence analysis. The input stimuli and output task were closely controlled. We demonstrated a strikingly similar bilateral neural network for both types of analysis. The network is right lateralised and includes the cerebellum, posterior superior temporal cortices, and inferior frontal cortices. These data are consistent with a common initial mechanism for the analysis of pitch and duration patterns within sequences. *NeuroReport* **10:3825–3830 © 1999 Lippincott Williams & Wilkins.**

Key words: Auditory; Cortex; Functional imaging; Human; Perception; PET; Sound sequence

Introduction

This study examines the brain mechanisms for the analysis of patterns of pitch and duration patterns in segmented sound. Such early processing is relevant to the perception of speech, environmental sounds and music [1], and is necessary for both human and non-human species.

Most human studies of sound sequence analysis have tended to use musical stimuli that are subject to culturally based rules [2]. Melody corresponds to a form of pitch sequence, and one form of rhythm corresponds to duration sequences. Studies of the analysis of melody and rhythm by the brain have employed a variety of techniques. In normal individuals, these include functional imaging [3–6] and psychophysical techniques [7]. Psychophysical testing of neurological patients with lesions is another technique that has demonstrated brain mechanisms for melody and rhythm analysis [8–10]. Another source of variation between studies is the degree of musical sophistication of the subjects, a factor likely to effect the strategy and neural basis for musical processing [6,7]. Whether the response requires local or global aspects of the sequence to be analysed is an important aspect, in terms of the processing strategies used and the underlying neural substrate [11]. The output task has also tended to differ between studies, both in terms of the amount of cognitive processing in addition to basic sequence analysis, and in terms of the mechanism of response.

Studies of the brain processing of melody or rhythm have produced conflicting results. For melody, different studies have suggested predominant processing in either the right [8,12] or the left [13] cerebral hemisphere. For rhythm, different studies have also suggested predominant processing in either the right [14,15] or the left [12,16,17] hemisphere. A role for the cerebellum in rhythmic analysis is suggested by several studies (reviewed in [5]). Differences in the suggested brain mechanisms for melody and rhythm processing might reflect differences in the techniques, subjects, stimuli, or tasks used. In particular, many melody studies have used musically sophisticated tasks requiring processing at a higher level than the analysis of basic sequence features (e.g. [12]), whilst rhythm tasks often require a patterned motor output corresponding to the stimulus pattern (e.g. [5,14,15]).

In this study we used PET to assess the brain activity during the analysis of sound sequences by musically naive subjects. We used sequences made up of elements with randomised pitch and duration. The output task is similar for both pitch and dura-

tion and does not depend on any musical knowledge or exposure. The approach is therefore designed to allow inference about the analysis of segmented sound below the level at which the sounds acquire symbolic significance. The stimuli for the pitch and duration sequence tasks and the final motor responses are the same, and the difficulty of the tasks is controlled. This experiment therefore allows an examination of differences between early processing of duration and pitch patterns in sequences.

Materials and Methods

Nine male subjects aged 21–31 years took part in the experiment, which was conducted with the approval of the ethical committee of the National Hospital for Neurology and Neurosurgery (London). All were strongly right-hand dominant [18]. The subjects were neurologically normal with no history of index neurological events or otolaryngological symptoms. A structured interview was used to assess musical listening experience, academic music lessons, practical music lessons and musical performance; no subjects had passed any musical grade exams (which was an exclusion criterion). Subjects received <1 h of training before scanning to ensure they understood the task.

Stimuli were six-element sequences (Fig. 1). The elements were all pure tones of 150 ms or 300 ms duration with 20 ms linear onset/offset ramps. The gap between each tone was always 50 ms. The first and last tone in the sequences was always identical. Sequences were created in pairs. For the first sequence of each pair the pitch of each tone was picked at random from five pitch values, chosen randomly from six groups of five pitches. The lowest pitch used in any sequence was A_5 (213 Hz) and the highest pitch $G^{\#}_5$ (819 Hz) and the groups of pitches were chosen so that the randomization could not generate any major intervals. The duration of each tone was also allocated randomly. For the second sequence of each pair the same pitch values were used for the elements, either in the same order or in a new random order (except for the first and last tones). The same duration values were used for the individual tones in the second sequence in the same order or in a new random order (except for the first and last tones). The pairs of sequences were thus of four possible types; same pitch/same duration, same pitch/different duration, different pitch/same duration and different pitch/different duration (Fig. 1). Four different sets of sequence pairs were created in this way and each used twice in the experiment.

The sequence sets were played to the subjects binaurally during scanning at a sensation level of 60 dB. Subjects were told to keep their eyes closed during presentation of the sounds and to attend to either the pitch or duration pattern within the sequences. After each pair of sequences, they were required to make a same–different response based on the attended dimension. Subjects were required to push one of two buttons with the right thumb after a high frequency tone. The pitch task and duration tasks were carried out for each of the four sets of sequence pairs. The tasks were carried out for $\sim$2 min during 80 s of which scan data were acquired.

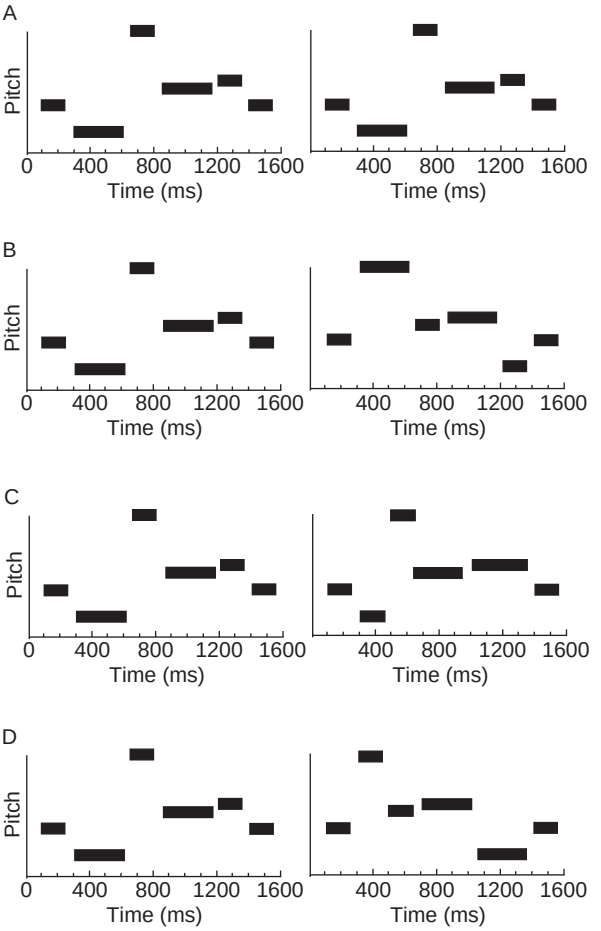

FIG. 1. Stimuli used. Examples of pairs of stimuli showing the following properties: (**A**) same pitch/same duration; (**B**) different pitch/same duration; (**C**) same pitch/different duration; (**D**) different pitch/different duration.

Subjects underwent 12 PET scans each using the ^{15}O-labelled water bolus technique and a Siemens scanner in 3D mode. Each subject underwent four scans when there was silence and no task, and eight scans during which the same four sequence sets were presented twice. The orders of stimuli and tasks were counterbalanced across subjects. Analysis was carried out using statistical parametric mapping software (SPM99; http//:www.fil.ion.ucl.ac.uk). Scans

were realigned and spatially normalized [19] to the standard sterotactic space of Talairach [20]. The data were smoothed with a Gaussian filter (full width at half maximum of 12 mm). Analysis of covariance was used to correct for differences in global blood flow between the scans. Differences in blood flow were assessed with the *t* statistic at each voxel.

Results

Table 1 shows the performance levels of the nine subjects overall and for the individual trials. All subjects achieved an overall performance for the two tasks at above chance level with the exception of subject nine for the pitch task. During debriefing, subjects reported the tasks to be difficult and to require constant vigilance. This is reflected in the scores shown in Table 1, with no subject achieving a 100% performance level for either task. Comparison of the performance of the pitch and duration tasks demonstrated no difference either overall (Wilcoxon signed ranks, $Z = -1.4$, $p > 0.10$) or for the individual sequence sets ($p > 0.05$ for all comparisons). Subjects did not experience the sequences as music. Performance of the subjects for each of the trials was included in the analysis as a covariate of interest.

Areas of activation for the contrast between the pitch analysis task and rest are listed in Table 2. Figure 2 shows the data rendered onto sections of the average T1 structural MRI for the nine subjects. This contrast with rest demonstrated significant activation in auditory and motor areas, for which there were *a priori* hypotheses. The region corresponding to the primary auditory cortex [21] was activated in the superior temporal planum bilaterally. These activations were part of large clusters of superior posterior temporal lobe activation that were significant at the $p < 0.05$ level (corrected for multiple comparisons); the activation extended posteriorly onto the planum temporale, lateral surface of the superior temporal gyrus and angular gyrus. Activation was also shown in the inferior colliculus. The left precentral sulcus also showed activation corresponding to the six button pushes with the right thumb during each scan ($Z = 4.15$, coordinates $-56,-8,50$).

Apart from the right and left temporal lobe clusters, significant activation of clusters at the $p < 0.05$ (corrected) level for the pitch task minus rest contrast was demonstrated in the cerebellum, right and left posterior inferior frontal cortex, anterior cingulate and the right dorsolateral prefrontal cortex. The cerebellar cluster involved both lobes of the cerebellum and the midline vermis in a symmetrical pattern. The cerebral activation in the network of significant clusters showed similar activation levels for corresponding areas on the two sides but greater spatial extent on the right. Activation that was not significant at the corrected cluster level occurred in the superior parietal cortex bilaterally; this activation was significant at the $p < 0.001$ (uncorrected) voxel level.

Areas of activation for the contrast between the duration analysis task and rest are listed in Table 3 and shown in Fig. 2. This shows a strikingly similar pattern of activation to the contrast between pitch and rest. Carrying out the contrasts pitch task minus duration task and duration task minus pitch task directly tested the presence of differences between the pitch and duration tasks. These contrasts

Table 1. Psychophysical performance of subjects

Subject	Task	Set 1 (score/10)	Set 2 (score/10)	Set 3 (score/10)	Set 4 (score/10)	Overall (score/40)
1	Pitch	7	5	8	5	25
	Duration	5	5	9	5	24
2	Pitch	7	8	6	4	25
	Duration	6	6	8	9	29
3	Pitch	8	6	8	5	27
	Duration	8	8	10	9	35
4	Pitch	10	9	7	8	34
	Duration	10	10	9	10	39
5	Pitch	8	5	6	7	26
	Duration	6	8	7	7	28
6	Pitch	10	5	8	5	28
	Duration	7	7	8	6	28
7	Pitch	10	10	9	10	39
	Duration	6	7	8	5	26
8	Pitch	7	6	10	6	29
	Duration	6	10	10	9	35
9	Pitch	4	5	7	2	18
	Duration	9	8	7	9	33

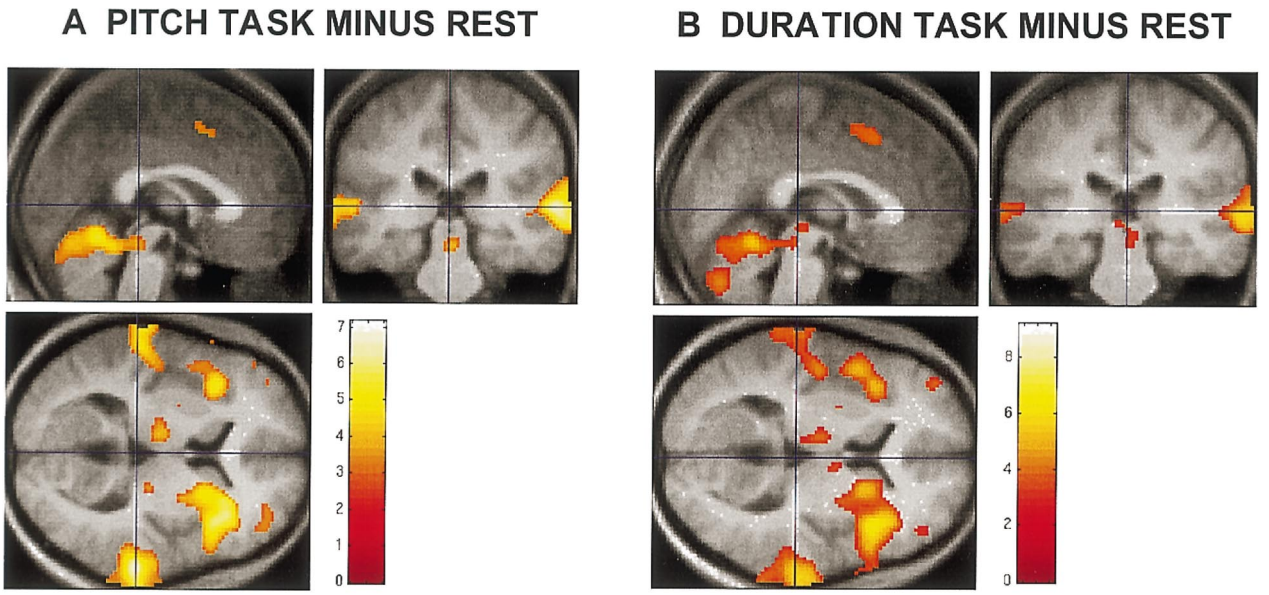

FIG. 2. Areas of significant activation at the level $p < 0.001$ (uncorrected) for (**A**) the pitch task minus rest contrast and (**B**) the duration task minus rest contrast. The areas are shown rendered onto sections of the mean T1 structural image for the nine subjects. Coronal sections are taken through the point with Talairach coordinates (A) 0,−30,6 and (B) 0,−26,6. The coronal and sagittal sections show activation within the inferior colliculus not listed separately in Table 1; this activation is contiguous with the cluster involving the cerebellum and cerebellar preduncle. Colour scale refers to z score.

Table 2. Pitch task minus rest contrast

Region	Coordinates (mm)			z score
	x	y	z	
Cerebellum	**−38**	**−60**	**−32**	**6.31**
	−18	**−66**	**−24**	**6.14**
	24	**−56**	**−24**	**5.98**
	−20	**−70**	**−30**	**5.98**
	40	**−64**	**−28**	**5.66**
	0	**−56**	**−12**	**5.23**
R posterior inferior frontal	**30**	**26**	**2**	**5.96**
	40	**18**	**8**	**5.67**
	26	**0**	**−2**	**5.42**
L posterior inferior frontal	**−36**	**20**	**0**	**5.93**
	−50	**2**	**−4**	**3.82**
	−48	**10**	**6**	**3.38**
R superior temporal	**66**	**−30**	**6**	**5.93**
	72	**−16**	**10**	**3.70**
	48	**−32**	**2**	**3.41**
R dorsolateral prefrontal	**46**	**38**	**20**	**5.67**
	30	**50**	**12**	**3.99**
Anterior cingulate	**10**	**26**	**36**	**4.69**
	−8	**18**	**42**	**4.50**
	8	**18**	**46**	**4.06**
L superior temporal	**−70**	**−28**	**4**	**4.66**
	−60	**−28**	**6**	**4.62**
L dorsolateral prefrontal	−38	50	18	4.61
R parietal	48	−64	54	4.49
L parietal	−50	−52	54	4.42

The clusters showing significant activation at the $p < 0.05$ level (corrected for multiple comparisons) are shown in bold. Other values are voxel-level activation at a significance level of $p < 0.001$ (uncorrected).
The three most significant maxima within each cluster are shown, except for the cerebellum, where six are shown.

Table 3. Duration task minus rest contrast

Region	Coordinates (mm)			z score
	x	y	z	
Cerebellum	**−16**	**−64**	**−22**	**7.63**
	−26	**−60**	**−30**	**5.87**
	−38	**−60**	**−28**	**5.83**
	38	**−64**	**−26**	**5.64**
	10	**−62**	**−18**	**5.36**
	24	**−68**	**−24**	**5.05**
R posterior inferior frontal	**42**	**18**	**8**	**6.07**
	30	**26**	**−8**	**5.12**
	50	**10**	**32**	**5.11**
R superior temporal	**68**	**−26**	**2**	**5.77**
	68	**−38**	**22**	**4.61**
	72	**−46**	**4**	**3.95**
R orbitofrontal	**30**	**52**	**−16**	**5.44**
R parietal	**50**	**−64**	**52**	**5.32**
L posterior inferior frontal	**−38**	**18**	**2**	**5.22**
	−48	**6**	**2**	**5.01**
L superior temporal	**−66**	**−38**	**12**	**4.36**
	−46	**−14**	**2**	**4.33**
	−66	**−26**	**6**	**3.89**
R dorsolateral prefrontal	**48**	**38**	**20**	**4.81**
	30	**50**	**12**	**3.79**
L dorsolateral prefrontal	**−40**	**50**	**16**	**4.27**
	−38	**52**	**6**	**3.68**
L orbitofrontal	−20	52	−18	4.71
L parietal	−48	−52	56	3.81

revealed no significant activation at the $p < 0.05$ (corrected) cluster level of significance.

Discussion

In this study we have demonstrated brain networks involved in the analysis of pitch and duration patterns in segmented sound. We have controlled subject inclusion, stimuli, task and task difficulty to allow inference about the processing of these patterns in all types of segmented sounds by normal listeners. We have demonstrated a bilateral network activated during both tasks that involves the cerebellum, posterior temporal lobes and posterior inferior frontal lobes, with right lateralization for the cerebral areas. The most significant activation was demonstrated in the cerebellum for both the pitch and duration tasks. A previous imaging study demonstrated cerebellar activation in rhythm tasks with a motor output following the stimulus [5]. The activation in the present study occured with an infrequent motor output that was not directly related to the actual stimulus pattern, for both pitch and duration tasks; this is in accord with a cerebellar role for sequence analysis in general, rather than a specific motor preparation in response to rhythm. The right-lateralized fronto-temporal activation shown for the pitch and rhythm tasks shows similarities to that demonstrated in a previous imaging study of melody perception [3]. This is consistent with such a network subserving the analysis of temporal pattern in sequences in general, rather than melody *per se*. The greater lateralisation in the previous study was striking and may reflect the more musical task used. Another previous imaging study suggested the involvement of both anterior and posterior temporal lobe areas in pitch sequence analysis [22]; the current study has demonstrated similar activation of the planum temporale during pitch sequence analysis but not activation of the anterior temporal lobe.

Conclusion

These data demonstrate a network for the processing of pitch and duration sequence processing. We have not disproved the null hypothesis that the processing of pitch and duration sequence is subserved by a common mechanism. Most models of pitch and rhythm analysis consider these processes in isolation. We suggest an exploration of models where pitch and duration sequences are analysed by common mechanisms.

References

1. Griffiths TD, Rees A and Green GGR. *Neurocase* (in press).
2. Dowling WJ and Harwood DL. *Music and Cognition.* London: Academic Press, 1985.
3. Zatorre RJ, Evans AC, and Meyer E. *J Neurosci* **14**, 1908–1919 (1994).
4. Zatorre RJ, Halpern AR, Perry DW *et al. J Cogn Neurosci* **8**, 29–46 (1996).
5. Penhune VB, Zatorre RJ and Evans AC. *J Cogn Neurosci* **10**, 752–765 (1998).

6. Schlaug G, Martin B, Thangaraj V *et al. Neuroimage* **3**, S318 (1996).
7. Bever TG and Chiarello RJ. *Science* **185**, 537–539 (1974).
8. Milner B. In: Mountcastle VB, ed. *Interhemispheric Relations and Cerebral Dominance*. Baltimore: Johns Hopkins University Press, 1962: 177–195.
9. Zatorre RJ and Samson S. *Brain* **114**, 2403–2417 (1991).
10. Liegeois-Chauvel C, Peretz I, Babai M *et al. Brain* **121**, 1853–1867 (1998).
11. Peretz I. *Brain* **113**, 1185–1205 (1990).
12. Shapiro BE, Grossman M and Gardner H. *Neuropsychologia* **19**, 161–169 (1981).
13. Gordon HW. *Cortex* **14**, 58–70 (1978).
14. Fries W and Swihart AA. *Neuropsychologia* **28**, 1317–1323 (1990).
15. Penhune VB, Zatorre RJ and Feindel WH. *Neuropsychologia* **37**, 315 (1998).
16. Robinson GM and Solomon DJ. *J Exp Psychol* **102**, 508–511 (1973).
17. Platel H, Price C, Baron J-C *et al. Brain* **120**, 229–243 (1997).
18. Elias LJ and Bryden MP. *Neuropsychologia* **36**, 37–43 (1998).
19. Friston K J, Ashburner J, Frith CD *et al. Hum Brain Mapp* **2**, 165–169 (1995).
20. Talairach P and Tournoux J. *A Stereotactic Co-planar Atlas of the Human Brain*. Stuttgart: Thieme, 1988.
21. Penhune VB, Zatorre RJ, MacDonald JD and Evans AC. *Cerebr Cortex* **6**, 661–672 (1996).
22. Griffiths TD, Buechel C, Frackowiak RSJ and Patterson RH. *Nature Neurosci* **1**, 421–427 (1998).

ACKNOWLEDGEMENTS: This work is funded by the Wellcome Trust (UK). We thank Richard Frackowiak for helpful comments.

Received 5 August 1999; accepted 30 September 1999