Deficits of musical timbre perception after unilateral temporal-lobe lesion revealed with multidimensional scaling

Séverine Samson,[1] Robert J. Zatorre[2] and James O. Ramsay[3]

[1]*Department of Psychology, URECA, University of Lille 3, France,* [2]*Montreal Neurological Institute, McGill University and* [3]*Department of Psychology, McGill University, Canada*

Correspondence to: Séverine Samson, Université de Lille 3 UFR de Psychologie – URECA, BP 149-59653 Villeneuve d'Ascq cedex, France
E-mail: samson@univ-lille3.fr

Summary

Thirty patients with unilateral temporal lobe excisions and 15 normal control subjects were tested in a task involving judgements of timbre dissimilarity in single tone and melodic conditions. Perceptual correlates of spectral and temporal parameters resulting from changing the number of harmonics and rise-time duration, respectively, were investigated by using a multidimensional scaling technique. The results of subjects with left temporal lobe lesion suggest that they were able to use the spectral and temporal envelopes of tones independently in making perceptual judgements of single tones. In the melodic condition, their results were significantly different from those of normal control subjects, suggesting that left temporal lesions do affect subtle aspects of timbre perception, despite these patients' preserved ability to make discrimination judgements using traditional paradigms. The major finding of this study concerns perceptual ratings obtained by subjects with right temporal lobe lesion, which revealed a disturbed perceptual space in both conditions. The most distorted results were obtained with single tones, in which the temporal parameter was less prominent. Tones were grouped according to their spectral content, but the results did not reflect a coherent underlying perceptual dimension. In general, the data from both patient groups (left lesions and right lesions) showed that the extraction of temporal cues was easier in the melodic than in the single tone condition, suggesting that the different durations and frequencies heard in a musical phrase enhance the importance of certain physical parameters. The findings of the present study replicate and extend previous results showing that timbre perception depends mainly upon the integrity of right neocortical structures, although a contribution of left temporal regions is also apparent. These data also demonstrate that multidimensional techniques are sensitive to more subtle perceptual disturbances that may not be revealed by discrimination paradigms.

Keywords: timbre; music; temporal lobe; perception

Abbreviations: LT = patients with left anterior temporal lobectomy; NC = normal control subjects; RT = patients with right anterior temporal lobectomy

Introduction

Musical timbre refers to the auditory parameters that allow one to distinguish musical instruments when frequency, intensity and duration remain identical (American Standards Association, 1960). For more than a century, musical timbre was associated almost exclusively with the spectral energy distribution of a tone (Helmoltz, 1868). However, the limits of such a definition were rapidly raised in the literature and several experimental findings suggested that other physical dimensions, such as the amplitude and phase patterns of components and particularly the temporal characteristics of tone, may influence timbre perception as well (Berger, 1964;

Schouten, 1968; Balzano, 1986). There is currently a consensus that musical timbre is a multidimensional property of sound (Plomp, 1970). The principal determinants of our perception of timbre are provided by the spectral energy and temporal variation of musical tones, as suggested by results of several psychoacoustical investigations (Miller and Carterette, 1975; Grey, 1977; Grey and Gordon, 1978; Wessel, 1979). However, the cerebral substrate associated with these acoustical attributes characterizing musical timbre has been investigated in very few studies (Milner, 1962; Tramo and Gazzaniga, 1989; Samson and Zatorre, 1994).

Neuropsychological results, as well as brain imaging, electrophysiological and anatomical findings suggest that neural systems in the human superior temporal gyrus are important in various aspects of auditory processing (Celesia *et al.*, 1968; Galaburda and Sanides, 1980; Zatorre *et al.*, 1992; Pandya, 1995; De Ribaupierre, 1997; Belin *et al.*, 2000). Therefore, these cerebral structures might also contribute to timbre perception. In 1962, Milner examined the effect of temporal lobe removal in a timbre discrimination task (Seashore *et al.*, 1940) in which the spectral information of harmonic tones was manipulated. Patients who had undergone a right temporal lobectomy for the relief of intractable epilepsy demonstrated a significant deficit compared with patients with left temporal lobectomy, suggesting that perception of spectral patterns relies on the integrity of the right temporal lobe. Similarly, results from commissurotomized patients showed that the right hemisphere was superior to the left on this same task, providing additional evidence in favour of the right hemisphere predominance for the processing of spectral information (Tramo and Gazzaniga, 1989). As opposed to the previously mentioned studies that focused on the processing of spectral information only, we recently examined timbre discrimination in patients with unilateral temporal lobe lesion by independently manipulating the temporal as well as the spectral envelopes of sounds (Samson and Zatorre, 1994). The results showed that right but not left temporal lobectomy impaired the discrimination of tones differing either by the number of harmonics (i.e. equivalent in this task to spectral envelope) or by the rise time (i.e. temporal envelope), suggesting that the right temporal lobe is essential for the processing of spectral as well as temporal information involved in musical timbre.

The predominant contribution of the right temporal lobe in the processing of spectral information is supported by other data reported in the literature, dealing with pitch processing. By using pitch perception tasks that mainly require integration of spectral information, it has been consistently found that this ability depends upon right hemisphere structures (Sidtis and Volpe, 1988; Zatorre, 1988; Divenyi and Robinson, 1989; Robin *et al.*, 1990). However, the predominant involvement of the right temporal lobe in processing temporal information seems to be in disagreement with several neuropsychological results. In numerous studies, deficits in the perception of time-related stimuli were reported in patients with left hemisphere lesions, suggesting that left hemisphere structures are predominantly involved in processing temporal information (Efron, 1963; Carmon and Nachshon, 1971; Swisher and Hirsh, 1972; Tallal and Newcombe, 1978; Robin *et al.*, 1990; Ehrlé *et al.*, 2001; Samson *et al.*, 2001). The apparent contradiction between these latter findings and the results obtained in timbre discrimination tasks seems to be attributed to the type of temporal factor manipulated in the tasks. Whereas the superiority of left hemisphere structures in temporal processing has generally been demonstrated when the temporal manipulation involves rapid temporal coding within a tens of milliseconds range presented sequentially (Tallal *et al.*, 1993; Belin *et al.*, 1998; Liégeois-Chauvel *et al.*, 1999; Ehrlé *et al.*, 2001; Samson *et al.*, 2001), the predominant involvement of the right temporal lobe in timbre processing has been reported when the modification is based on slower temporal changes, within the hundreds of milliseconds range (Samson and Zatorre, 1994). This leads us to conclude that the processing of spectral energy and temporal variations involved in musical timbre depends on the integrity of the right temporal neocortex.

In all previously reported studies, the discrimination paradigm used involves perceptual judgements of same and different trials that vary along one physical dimension at a time. For instance, in the timbre discrimination task that we designed in a previous study, we separately manipulated either the spectral or the temporal envelopes of sound by changing the number of harmonics or the rise-time duration (Samson and Zatorre, 1994). As already emphasized, musical timbre does not depend upon a single physical attribute, but should be considered as a multidimensional property of sound. It seems, therefore, more relevant to use paradigms that allow investigation of multidimensional stimuli that vary along several dimensions at the same time. An especially well-suited candidate for examining such stimuli is the multidimensional scaling technique (Shepard, 1962; Kruskall, 1964). Multidimensional scaling attempts to represent the dissimilarities between a set of physically different stimuli as distances between points in an *n*-dimensional space (Ramsay, 1982*a*). For this purpose, trials composed of different stimuli are presented to subjects to judge their similarity on a rating scale. By using matrices of perceptual measures, an algorithm operates to maximize a goodness-of-fit function relating the distances between the points to the corresponding dissimilarity ratings between the sounds. This approach provides a powerful tool for the representation of stimulus–attribute relationships but has rarely been adapted to neuropsychology. Moreover, it allows study of the extent to which physical distances between multidimensional acoustic information are actually perceived by a subject, instead of investigating only the ability to distinguish same from different pairs of tones. As emphasized by Whitfield (1985), such similarity judgements based on the extraction of particular features play an important role in recognition by requiring higher order perception than discrimination or detection of differences.

In the musical domain, multidimensional scaling methods have been used in normal listeners to better understand the perceptual dimensions underlying the processing of musical timbre. Several studies investigated the dependence of timbre on the spectral amplitude pattern (Plomp, 1970; Wessel, 1973; Miller and Carterette, 1975; Gordon and Grey, 1978; Grey and Gordon, 1978). However, few studies have conclusively demonstrated the importance of the temporal envelope in timbre perception (Grey, 1977; Wessel, 1979; Krimphoff *et al.*, 1994; McAdams and Cunible, 1992; McAdams *et al.*, 1995), suggesting that this acoustical
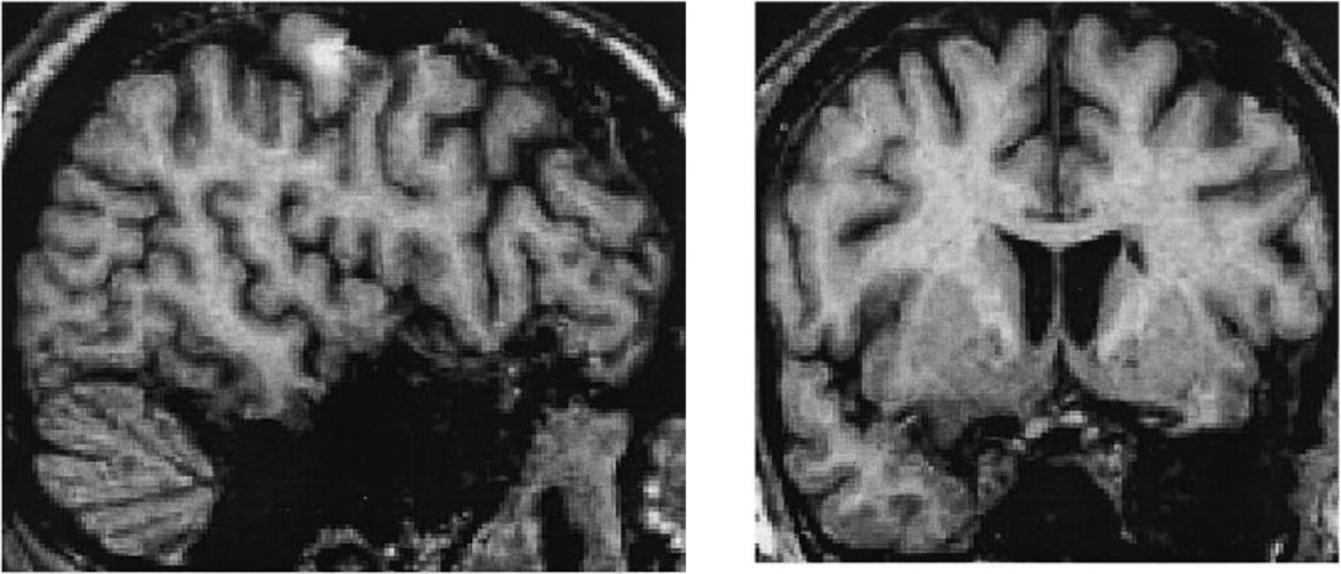
Fig. 1 MRI of a representative right temporal lobectomy (*left*: sagittal section; *right*: coronal section).

attribute is a less salient feature than the spectral characteristic. In a recent study (Samson *et al.*, 1997), we succeeded in demonstrating the dependence of timbre on both spectral and temporal envelopes of tones. By manipulating the spectral and temporal characteristics of synthetic tones, we were able to show that musically unselected listeners could recover in their perceptual judgements the unambiguous acoustic dimensions that were physically inherent to the stimuli. Specifically, we found that number of harmonics and rise time were reflected as orthogonal dimensions in a two-dimensional perceptual array, suggesting that this information can be perceived in a gradual and ordered way by normal listeners, who are able to use such information to make perceptual judgements of musical timbre independently. This paradigm validated in normal subjects offers an appropriate tool for neuropsychological purposes, providing the opportunity to differently investigate the nature of dysfunctional perceptual abilities in brain-damaged subjects.

Another issue raised in this latter study (Samson *et al.*, 1997) concerns the different processing involved in the perception of isolated tones compared with melodic sequences. Most studies reported in the literature have investigated timbre perception using single tones, neglecting the perception of sequences of tones. However, real music usually involves patterns of different notes played over time, in which a more complete picture of a given timbre might emerge. Throughout a musical phrase, different durations and frequencies are heard, and the physical parameters of tones evolve. This evolution could obscure or enhance the importance of certain parameters. It therefore seems ecologically valid to examine timbre perception in melodies as well as in single tones, and to compare the spatial configurations resulting from these two types of stimuli. Although we

obtained similar results with both melodies and single tones in our previous study carried out in normal participants, we can predict that the timbre perception of these two different types of stimuli might be differently affected by damage to the temporal neocortex.

The goal of the present study was to further explore the role of each temporal lobe in timbre perception using the previously described paradigm based on multidimensional scaling analysis of dissimilarity judgements (Samson *et al.*, 1997). To demonstrate how the organization of the perceptual space produced by multidimensional scaling analysis might depend upon the integrity of the temporal neocortex, we tested patients who had undergone a right or a left temporal lobectomy for the relief of medically intractable epilepsy and compared their performance with results of normal control subjects. The subjects included in the present investigation also participated in a previously published study examining timbre perception with a conventional discrimination tool (Samson and Zatorre, 1994), which allowed direct comparison of the two perceptual procedures. According to the previously reported results, we hypothezised that patients with lesions of the right temporal lobe would show a distorted pattern of perceptual dissimilarity ratings, being unable to perceive spectral and temporal differences between the stimuli. No specific distortions were predicted for patients with lesions of the left temporal lobe, who are generally not impaired in timbre discrimination tasks. Since we tested timbre perception with melodies as well as single tones, we further predicted that the representation for the melodic stimuli would be more stable than the one obtained for the single tone stimuli, allowing the different groups of subjects to process more efficiently both spectral and temporal information.

Table 1 *Characteristics of subjects: sex distribution, mean age, and mean years of education as well as Full Scale IQ (range)*

Group	Sex		Age (years)	Education (years)	Wechsler IQ (FSIQ)
	F	M			
Normal control	9	6	26.2 (14–47)	14.4 (9–17)	–
Left temporal	9	6	28.2 (15–42)	14.0 (8–18)	96.9 (80–124)
Right temporal	7	8	30.0 (20–41)	13.3 (10–17)	96.1 (77–114)

Methods

Subjects

Thirty subjects who had undergone a left (LT; $n = 15$) or a right (RT; $n = 15$) anterior temporal lobectomy at the Montreal Neurological Institute for the relief of intractable partial complex seizures participated in this experiment, as well as 15 normal control (NC) subjects. All these subjects were tested in a previous study examining timbre discrimination (Samson and Zatorre, 1994). All gave informed consent to participate in this study, which was approved by the Research Ethics Committee of MNI.

Except when noted, the epileptogenic lesions dated from birth or early life and were static and atrophic as determined by pathological analysis. The temporal lobe excisions always included the temporal neocortex anterior to Heschl's gyrus, the temporal pole, the amygdala and various amounts of the hippocampus and parahippocampal gyrus. An illustration showing an MRI of an excision representative of those in the present study is displayed in Fig. 1.

The patients included in this study were presumed to have language function lateralized to the left cerebral hemisphere based on their right-handedness and on dichotic listening measures (Zatorre, 1989); this was confirmed by intracarotid Sodium Amytal studies (Wada and Rasmussen, 1960) in the case of three left-handed subjects. All subjects with known atypical language representation (bilateral or right hemisphere), or in whom there was radiological or clinical evidence of rapidly growing neoplasm, diffuse cerebral damage or bilateral independent electrographic abnormality, were excluded. Also excluded were patients with a full scale IQ of <75 on the WAIS-R (Wechsler Adult Intelligence Scale—Revised; Wechsler, 1987), or any known hearing loss.

Musical experience, as determined by responses to a brief questionnaire, was balanced among the groups of subjects. Seventy-five per cent of the subjects had little or no formal musical experience, and no professional musician participated in this study.

Table 1 presents the sex distribution and years of education for each group, as well as WAIS-R full scale IQ values for the patient groups. Separate one-way ANOVAs (analyses of variance) indicated no significant difference between the subject groups with respect to age [$F(2,42) = 0.96$; $P = 0.39$] or educational level [$F(1,28) = 0.02$; $P = 0.85$]. There was also no significant difference between the patient groups in full scale IQ [$F(1,28) = 0.02$; $P = 0.85$].

Left temporal lobe group

This group included one case of ganglioglioma and two cases of oligodendroglioma. Seven of the 15 patients had a large removal from the hippocampus and parahippocampal gyrus, while the remaining cases had a small resection of those structures. None of these patients had a removal involving the transverse gyri of Heschl. Five patients were tested 2 weeks post-operatively, whereas the remaining patients were tested during follow-up evaluation 1 year or more after surgery (mean 2.5 years).

Right temporal lobe group

This group included one case of oligodendroglioma. Nine of the 15 patients had a large removal from the hippocampus and parahippocampal gyrus, while the remaining cases had a small resection of those structures. Three patients had a partial or complete removal of the transverse gyri of Heschl. Six patients were tested 2 weeks post-operatively, whereas the remaining patients were tested during follow-up studies (mean 2.6 years).

Normal control group

Fifteen normal right-handed subjects without known neurological problems were also tested. These subjects, drawn from hospital staff and patients' families, were matched to the patient groups with respect to sex, age, educational level and musical experience.

Stimuli

The stimulus sounds were synthesized by selecting three levels of spectral change and three levels of temporal change. The spectral change corresponded to the number of harmonics: either one, four or eight harmonics. The temporal change consisted of the manipulation of the linear rise time, which could be 1, 100 or 190 ms long, immediately followed by a linear decay time, without a steady-state portion. These

parameters were selected on the basis of pilot data indicating that the stimuli would be clearly distinct from one another. The minimum and maximum rise times (1 and 190 ms) were chosen as a function of the total duration of the shortest tone to be used in the study (200 ms). Representations of the wave form and spectra of the stimuli are displayed in Fig. 2. Nine hybrid sounds (resulting from crossing the three levels of spectral composition with the three levels of rise time) were then created by digital synthesis at a 16 kHz sampling rate. The stimuli did not sound like actual instruments, thereby avoiding any semantic level confound. The Mitsyn Command Language (Henke, 1976) was used to synthesize the sounds that were stored on the hard disk of a Compaq 386 computer. Stimuli were presented binaurally to each subject through Sennheiser 222 headphones after low-pass filtering at 8 kHz.

Single tones

A 400 ms tone played on the musical tone G4 (fundamental of 396 Hz) was created for each timbre, therefore making nine tones with the same duration, fundamental frequency and intensity. Each trial was composed of two tones separated by a 500 ms silence. All the tones were presented at 70 dB SPL (A scale), as measured by a GenRad sound pressure meter at the headphone.

Melodies

An eight-note melody was synthesized using each of the nine musical timbres. This short melody involves a rhythmical pattern with three different tone durations, and spans an octave. To prepare these different tones, the rise time was kept constant for each type of temporal manipulation but the decay time was varied according to the tone duration. For instance, a rise time of 100 ms was followed by either 100, 300 or 500 ms for the three note durations of 200, 400 and 600 ms, respectively. These values were chosen on the basis of pilot studies, which indicated that a constant rise time elicited a stable percept, whereas a rise time proportional to the length of the tone yielded different percepts for each value. Each trial was composed of two melodies separated by 500 ms silence.

Procedure

The nine stimuli were presented to each subject in an all-pairs design with no stimulus paired with itself, which resulted in the presentation of 36 stimulus pairs. Two sets of 36 pairs were prepared, the second set consisting of the same pairs of stimuli presented in the reverse order. The identical procedure was used for the single tones and for the melodies. Therefore, we had two sets of 36 pairs of tones as well as two sets of 36 pairs of melodies.

Subjects were told that the aim of this study was to examine their ability to distinguish different types of sounds. The subject's task was to judge the dissimilarity between two

stimuli. On each trial, two stimuli (either tones or melodies) were played on different timbres and, immediately after their presentation, a question mark appeared on the computer screen. The subject responded with the keyboard by selecting a number corresponding to the degree of dissimilarity. The dissimilarity rating was made on scale of 1 ('very similar') to 8 ('very different'). The program waited until the subject gave his response before presenting the next trial, which occurred 2 s after the response.

The melodic and tonal versions of the test were presented in the same experimental session. All subjects started with one set of melodies (36 pairs) followed by one set of tones (36 pairs), followed by the other set of melodies (reverse order from the first) and then the other set of tones. This order of presentation was chosen because the melodic version was easier to perform than the tonal one. Half of the subjects started with the first set of melodies and the first set of tones while the other half started with the second set of melodies as well as the second set of tones. The order of presentation of each pair of stimuli within a set was completely randomized by the computer.

Training sessions

Because subjects were generally unsophisticated musically, they received training prior to the experimental tasks to ensure that they were able to use the rating scale adequately. Preliminary tasks requiring perceptual judgement of tones that differed by pitch and loudness were presented to each subject before the experimental tasks. They consisted of 20 examples followed by two series of 36 pairs of tones involving dissimilarity judgements on an eight-point scale. Then, the different timbres were exposed to each subject through 15 examples consisting of pairs of melodies played on different timbre for purposes of familiarization. Examples of the extremes were given by presenting a pair of two very different stimuli and a pair of two very similar ones, followed by 13 other samples of different pairs of melodies. Each pair of melodies was followed by a written message on the screen to comment the type of dissimilarity between the timbre (i.e. very different, very similar or quite different) and to illustrate the sort of differences the subject should expect to hear. Then, the experimental task *per se* was demonstrated to the subject through five examples consisting of pairs of melodies. The subject had to rate the dissimilarity between the two members of the pair on an eight point scale, without feedback, in a manner identical to the experimental task.

Results

The dissimilarity judgements for each subjects were stored as a 9×9 matrix of data minus the diagonal. The matrices of dissimilarities for all the subjects were then processed by the multidimensional scaling program MULTISCALE II, developed by Ramsay (1982*b*). Multidimensional scaling analyses were carried out separately for each group of subjects. The
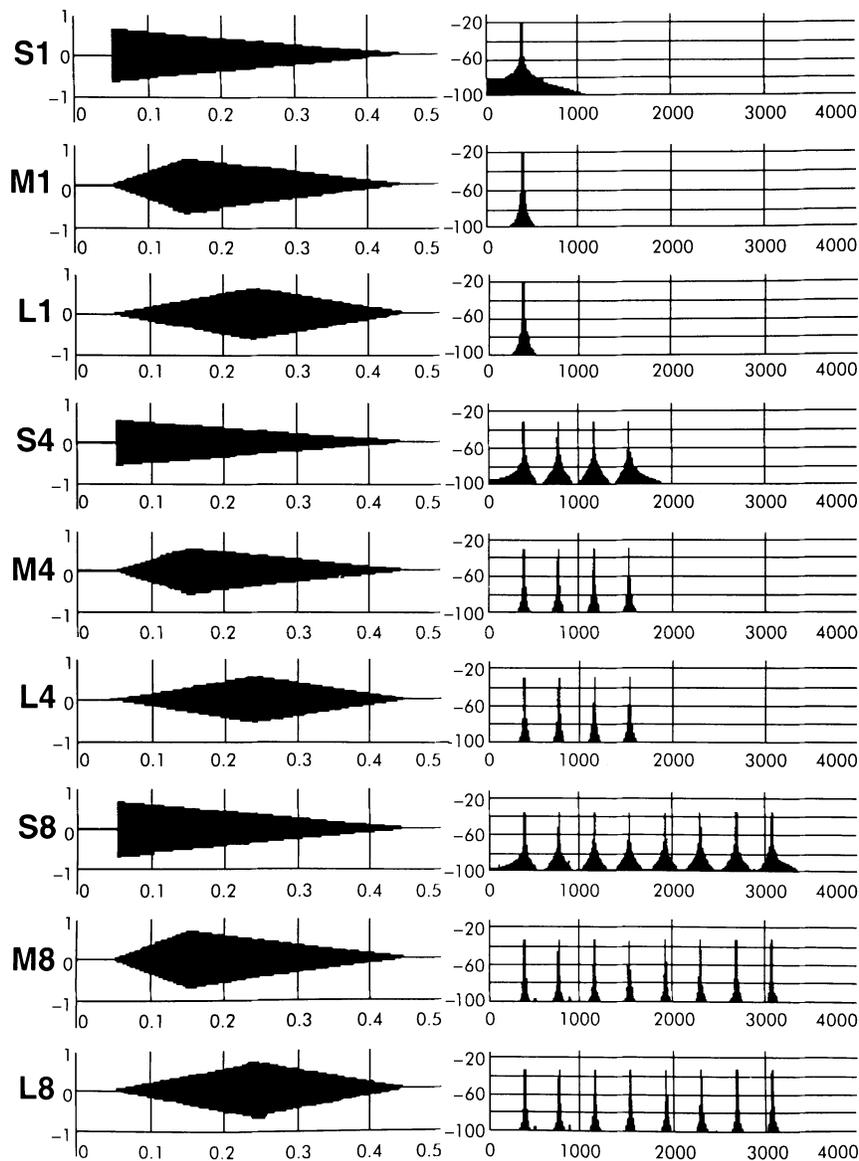
Fig. 2 Representations of each of the nine stimuli differing by spectral and temporal parameters. The left panel in each case shows the temporal waveform of the stimulus (abscissa: time, in tenths of seconds; ordinate: amplitude, arbitrary scale). The right panel in each case shows the Fourier spectrum of the stimulus (abscissa: frequency, in Hz; ordinate, energy level, in dB). S, M, L correspond to short (1 ms), medium (100 ms) and long (190 ms) rise times, respectively. Numbers 1, 4 and 8 refer to the number of harmonics present in the stimulus. Thus, for example, S1 corresponds to a stimulus with short onset (1 ms; note sharp attack on waveform in left panel) and consisting of a single harmonic (note the presence of only one harmonic in the right panel).

spatial configuration obtained from these analyses was the result of maximum likelihood estimation, and showed a good statistical relationship between the distances separating the points in the space and the dissimilarity judgements given for the pairs of stimuli. Preliminary analysis showed that there were few meaningful differences in judgement with respect to the order of presentation. We therefore combined the data obtained during the two orders of presentation used for the single tones as well as for the melodies by averaging the

responses for the two orders. The multidimensional scaling algorithms were used to generate geometric representations of the sounds in a Euclidean space, in which the points correspond to the nine different tones. Based on results of a previous study carried out in NC subjects, it was determined that a three-dimensional scaling solution was most appropriate for the data (Samson *et al.*, 1997). Therefore, spatial representations were obtained for three dimensions. The analyses of the single tones will be considered separately

from the analyses of melodies, and results of the three groups of subjects will be presented successively. Differences between groups were assessed by using χ^2 derived from the three-dimensional scaling solutions obtained for the two groups of subjects and for a combined group, including the subjects of each group involved in the comparison. More

Normal control group: single tones

Left temporal lobe group: single tones

Right temporal lobe group: single tones

specifically, the log likelihood ratio of the combined group was subtracted from the sum of the log likelihood ratios obtained by each group multiplied by two.

Single tone condition

The coordinates of the nine stimuli for all three dimensions resulting from the multidimensional scaling obtained by each group of subjects are presented in Table 2 for the single tone condition.

Normal control group

Three dimensions accounted for 98% of the variation among the points. However, it was evident that the first two dimensions were far more important than the third. The percentage of variance was evenly distributed between the first two dimensions (48 and 38%), whereas the third dimension only accounted for 12% of the variance. Figure 3 (top panel) shows Dimension 2 of the scaling solution for all normal subjects plotted against Dimension 1. The points corresponding to the sounds that are judged very dissimilar are distant, while the ones corresponding to the sounds that are judged very similar are close to each other.

The distribution of points in this plane is very systematic and is clearly related to the physical characteristics of sound. The two dimensions appear also to be orthogonal. Along the first dimension, the sounds were grouped according to the number of harmonics. The one-harmonic tones were separated from the four-harmonic tones by a distance similar to the distance separating the four-harmonic from the eight-harmonic tones. Along the second dimension, tones with a 1 ms rise time were separated from tones with 100 and 190 ms rise times quite well. The distance between points corresponding to tones with 1 and 100 ms rise time was larger than the distance between points corresponding to tones with 100 and 190 ms rise times. The third dimension distinguished the tones composed of four harmonics from those with one or eight harmonics (see Table 2).

Left temporal group

Three dimensions derived from multidimensional scaling accounted for 99% of the variation among the points. However, it was evident that the first dimension, which
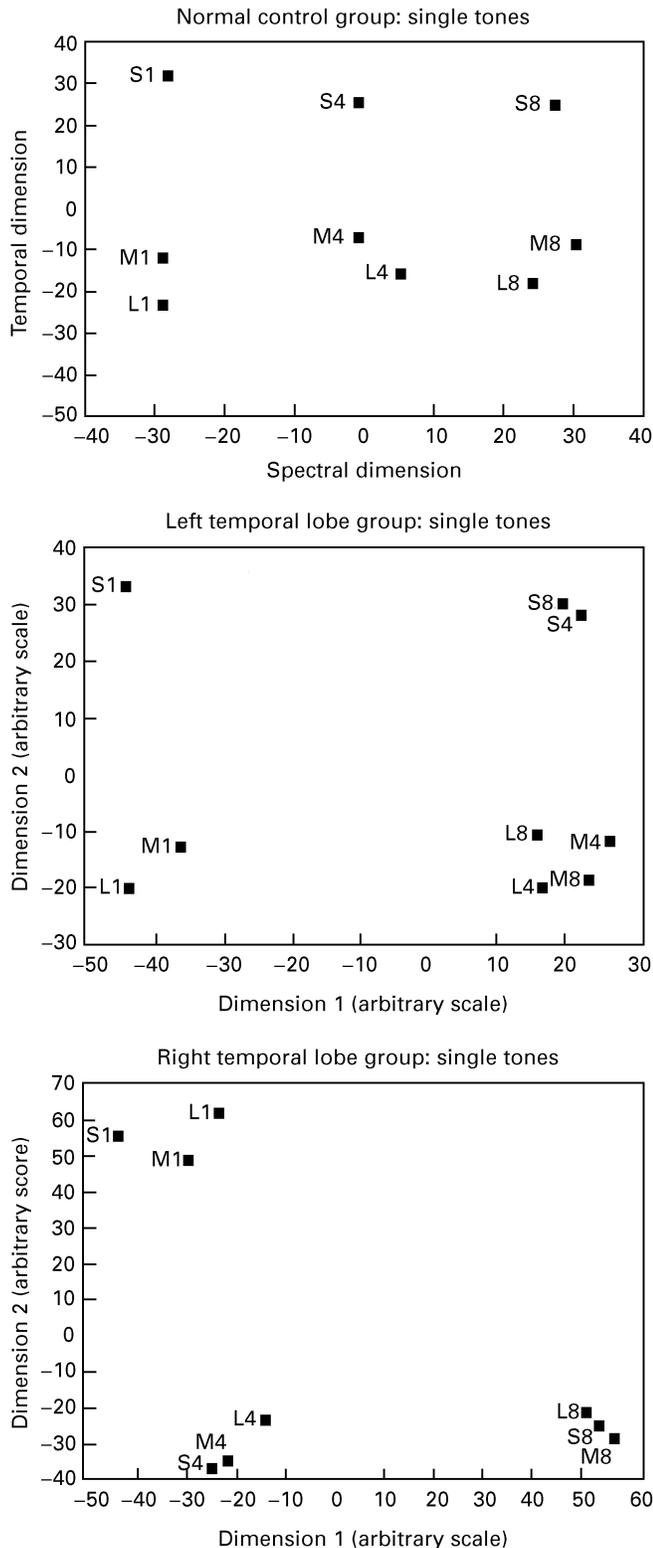
Fig. 3 Two-dimensional spatial representation based on the dissimilarity ratings for the nine single tones, as revealed by multidimensional scaling analysis (scales used for Dimensions 1 and 2 are arbitrary). Each point represents the position of one stimulus in the perceptual space; the labels (S1, S4, etc.) correspond to the stimuli, as described in Fig. 2. The configurations obtained by each group of subjects (NC = normal control; LT = left temporal; RT = right temporal) are represented separately.

Table 2 *Coordinates of each point in the spatial configuration for the single tone condition for the three groups of subjects*

Stimulus	No.	NC			LT			RT		
		Dimension			Dimension			Dimension		
		x 1	y 2	z 3	x 1	y 2	z 3	x 1	y 2	z 3
1S	1	−28	32	15	−45	34	−23	−44	56	38
1M	2	−29	−12	8	−37	−13	−23	−30	50	0
1L	3	−29	−23	2	−44	−20	−15	−24	63	−12
4S	4	−1	26	−15	22	29	−17	−25	−37	30
4M	5	−1	−7	−13	27	−12	−16	−22	−35	−25
4L	6	5	−15	−19	18	−20	−11	−15	−23	−14
8S	7	27	25	7	19	30	33	53	−24	33
8M	8	30	−9	5	24	−18	34	56	−28	−16
8L	9	24	−18	10	16	−10	39	51	−22	−33

Dimensions 1 and 2 correspond to the abscissa and the ordinate, respectively, of the graphical representation of the data shown in Fig. 3.

Table 3 *Coordinates of each point in the spatial configuration for the melodic condition for the three groups of subjects*

Stimulus	No.	NC			LT			RT		
		Dimension			Dimension			Dimension		
		x 1	y 2	z 3	x 1	y 2	z 3	x 1	y 2	z 3
1S	1	−29	41	5	−42	35	1	−59	45	−11
1M	2	−34	−8	15	−38	−9	8	−60	−16	−20
1L	3	−38	−17	5	−41	−20	2	−56	−38	−13
4S	4	12	36	7	9	39	17	23	52	−24
4M	5	9	−18	17	18	−20	18	26	−14	−28
4L	6	7	−29	10	9	−25	25	26	−36	−18
8S	7	25	34	−18	29	38	−19	22	47	41
8M	8	21	−14	−24	30	−13	−23	41	−8	38
8L	9	27	−25	−18	26	−24	−29	36	−33	34

Dimensions 1 and 2 correspond to the abscissa and the ordinate, respectively, of the graphical representation of the data shown in Fig. 4.

accounted for 44% of variance, was far more important than the second and the third dimensions (24 and 31%, respectively). The results for the LT group seemed less orderly than for the NC group, but the difference between the two configurations did not reach significance ($\chi^2 = 28.7$; $P > 0.05$).

As shown in Fig. 3 (middle panel), the first dimension allows to group the sounds according to the number of harmonics, but in this case the one-harmonic tones were separated and very distant from the four-harmonic and eight-harmonic tones. Along the second dimension, sounds were grouped according to their rise time and tones with a 1 ms rise time were clearly separated from tones with 100 and 190 ms rise times, the distance between points corresponding to tones with 100 and 190 ms rise times being very small. The third dimension distinguished the tones composed of eight harmonics from those with one and four harmonics (see Table 2).

Right temporal group

As for the other two groups, the three dimensions accounted for 98% of the variation among the points. The percentage of variance was evenly distributed between the first two dimensions (39 and 43% for first and second dimensions, respectively), whereas the third dimension only accounted for 16% of the variance. As expected, RT subjects displayed a significantly different pattern of results ($\chi^2 = 198$; $P < 0.001$) from the NC group.

The very distorted configuration displayed in Fig. 3 (bottom panel) showed that RT subjects group the sounds almost exclusively according to the number of harmonics, but in a very different way to the other two groups. Dimension 1 dissociates tones with eight harmonics from tones with one or four harmonics, whereas Dimension 2 distinguishes tones with one harmonic from tones with four or eight harmonics. The third dimension distinguished the tones with a 1 ms rise from those with 100 and 190 ms rise times (see Table 2).

Although RT subjects did tend to group according to harmonic structure, their perception of spectral characteristics of tones was not systematically ordered in the way that it was for NC and LT subjects. The distance between points corresponding to tones with 1, 100 and 190 ms rise times was very small, suggesting that these patients did not perceive differences between tones with different rise time. This result suggests that RT subjects neglected the time information, and relied almost exclusively on the spectral cues.

Melodic condition

The coordinates of the nine stimuli for all three dimensions obtained by each group of subjects are presented in Table 3 for the melodic condition.

Normal control group

Modelling in three dimensions accounted for 98% of the variance, the first two dimensions being clearly dominant (accounting for 46 and 39% of the variance, respectively) compared with the third one (13%). The spatial configuration of the nine different points is very similar to the one described for the single tones, as shown in Fig. 4 (top panel).

The distribution of the points in this configuration is very well ordered. The first dimension displayed on the horizontal axis showed three groups of points corresponding to timbres consisting of one, four and eight harmonics, equally distributed along this dimension. Dimension 2 separates points along the vertical axis that correspond to tones with different rise times. Again, the 1 ms rise time tones were separated from the 100 and 190 ms rise time tones, these two latter types of sounds being very close to each other. Finally, the third dimension, which accounted for only 13% of the variance, differentiated points corresponding to eight harmonic tones from points corresponding to one and four harmonic tones (see Table 3).

Left temporal group

Again, the three dimensions accounted for 99% of the variation among the points. The percentage of variance was evenly distributed between the first two dimensions (45 and 37% for the first and second dimensions, respectively), whereas the third dimension only accounted for 17% of the variance. As opposed to single tone condition, the solution obtained by the LT group in the melodic condition was significantly different from the result for the NC group ($\chi^2 = 61.9$; $P < 0.005$), suggesting that judgements provided by the LT subjects were really less stable than judgements given by NC subjects.

Figure 4 (middle panel) shows the geometrical representation of the first two dimensions derived from the scaling solution for LT subjects. The sounds were grouped according to the number of harmonics along the first dimension, but the one-harmonic tones being clearly separated from the four-

and eight-harmonic tones. Along the second dimension, the sounds were grouped according to the temporal information since tones with a 1 ms rise time were clearly distant from
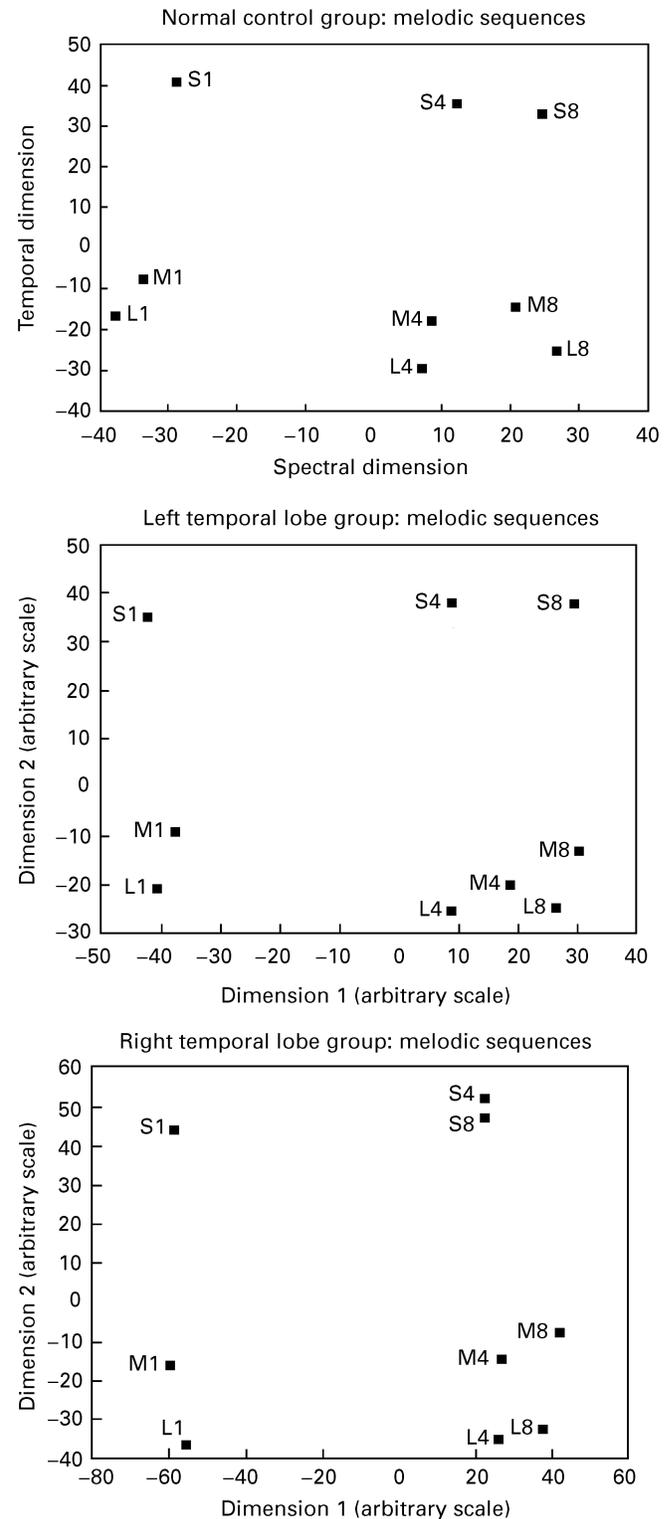
Fig. 4 Two-dimensional spatial representation based on the dissimilarity ratings for the nine melodic patterns, as revealed by multidimensional scaling analysis. Other details as in Fig. 3.

tones with 100 and 190 ms rise times. The distance between points corresponding to tones with 100 and 190 ms rise times was very small, suggesting once more that these patients really differentiate tones with 1 ms rise time from tones with longer rise times (100 and 190 ms rise time). The third dimension distinguished the tones according to their harmonic composition: the tones with one harmonic were separated from those with four harmonics, which were also distant from those with eight harmonics, but this dimension accounted for a small amount of variance (see Table 3).

Right temporal group

Again, the three dimensions accounted for 99% of the variation among the points. Dimension 1 accounted for almost half of the variance (46%), the remaining variance being distributed between the last two dimensions (33 and 20% for Dimensions 2 and 3, respectively). As predicted, the configuration of the RT group was distorted compared with the NC group, and this finding was corroborated by statistical analysis ($\chi^2 = 55.7$; $P < 0.05$).

Figure 4 (bottom panel) shows Dimension 2 of the scaling solution for RT subjects plotted against Dimension 1. Along the first dimension, which accounts for almost half of the variation, the one-harmonic tones were separated and very distant from the four- and eight-harmonic tones. Differently from the single tone condition, RT subjects did tend to group the tones according to their rise time. Along the second dimension, which accounts for one-third of the variance, tones with a 1 ms rise time were clearly separated from tones with 100 and 190 ms rise times, suggesting that when notes are played over time in a melody, the emergence of temporal cues involved in musical timbre seems to be more salient for RT subjects. The third dimension distinguished the tones composed of one or four harmonics from those with eight harmonics (see Table 3).

Discussion

The goal of the present study was to investigate perception of musical timbre by using a procedure based on the multi-dimensional scaling method. As opposed to the traditionally used discrimination paradigm, in which subjects differentiate same from different trials that vary along a single dimension, we were interested in exploring how patients with temporal lobe damage actually perceive similarities between timbres that vary along several dimensions. By analysing similarity judgements obtained in normal subjects as well as in patients who had undergone unilateral temporal lobe removal, we were able to describe qualitative differences between the three groups of subjects in processing the spectral and the temporal envelopes of a periodic complex tone.

Results obtained in normal participants indicate that they are implicitly sensitive to spectral and temporal information in musical tones, and are able to use them independently and orthogonally in making perceptual judgements of musical timbre in a gradual and ordered fashion for single tones as well as for melodies. Thus, even musically untrained listeners can systematically order stimuli in a perceptual space that reflects the acoustical structure of the stimuli. More importantly, the present data indicate that it is possible to demonstrate the independent contribution of temporal features to timbre perception, in addition to the well established contribution of harmonic structure. The geometrical model resulting from multidimensional scaling analysis was composed of three dimensions, but the third one was less salient and less clearly defined.

The first dimension displayed on the spatial configuration distinguishes points corresponding to tones with different numbers of harmonics, the one-, four- and eight-harmonic tones being clearly separated along this dimension. The second major dimension resulting from the multidimensional analysis was related to the temporal characteristic of the tones, which was determined by the rise/decay time. The distances separating tones with 1 and 100 ms rise times were consistently larger than the one separating tones with 100 and 190 ms rise times, suggesting that subjects were able to differentiate 1 from 100 ms rise time clearly, but that the perceptual difference between tones of 100 and 190 ms rise time was less salient. From the physical point of view, very abrupt onsets (e.g. 1 ms rise time) produce a certain amount of energy across the spectrum, producing an interaction between temporal and spectral cues (Fig. 1). The availability of both cues can explain why the perceptual distance separating tones with 1 ms rise time from the others is largest. It could also be attributed to the logarithmic representation of the temporal dimension, as suggested by several authors (Krimphoff *et al.*, 1994; McAdams *et al.*, 1995). However, the relative order of rise time (1 ms, followed by 100 ms, followed by 190 ms) was systematically preserved in both sets of stimuli, emphasizing the reliability of this temporal dimension. The third dimension accounted for only a relatively small amount of variance, and the ordering of stimuli was different for single tones and for melodies. However, it was related to the harmonic composition of tones in both conditions, indicating again the important contribution of spectral cues in timbre judgements. A condition involving melodic sequences has rarely been examined in past studies of timbre, although it appears to be a slightly more realistic situation than the single tone condition; it did not demonstrate a more orderly solution than the single tone condition, but differences did occur in patients with temporal lobe lesions, as discussed below.

The spatial configuration displayed by LT subjects suggests that they too are able to make use of these acoustical attributes in perception. However, they do not perceive discrete physical changes as systematically as NC subjects. Despite the apparent difference between the pattern of results obtained for the single tone condition by the LT subjects and by the NC subjects, the statistical analysis did not reveal a reliable effect. Indeed, the LT subjects were able to recover three independent perceptual dimensions. The first dimension, which is related to spectral information, allowed

separation of pure tones (one harmonic) from complex tones (four and eight harmonics). The second dimension had to do with the temporal characteristic of the tones, which was determined by the rise/decay time. Along this dimension, LT subjects can dissociate abrupt onsets (e.g. 1 ms rise time) from longer ones (e.g. 100 and 190 ms rise times). The third dimension also concerned the spectral characteristic of tones, allowing separation of the eight-harmonic tones from the one- and four-harmonic tones. The results from the LT subjects obtained for the single tones are in agreement with data reported in a previous study (Samson and Zatorre, 1994), in which we tested discrimination of spectral and temporal information involved in musical timbre. The absence of significant differences between performance of LT and NC subjects leads us to conclude that timbre perception does not depend to a great extent on the integrity of the left temporal neocortex, at least when isolated tones are used.

The spatial configuration obtained for the LT group in the melodic condition seems to provide a more orderly solution than the configuration obtained for the single tone condition. The first and the second dimensions became more salient in this condition, and seem to better dissociate the tones according to the number of harmonics and to the rise time, respectively. The additional information provided by the frequency excursions and different durations of the tones within the melodies may have facilitated access to the spectral and temporal features involved in musical timbre. However, the solution obtained by LT subjects in this condition was significantly different from the result of the NC subjects, suggesting that the representation for LT subjects was less stable than the representation for the NC subjects. The sequential presentation of tones characterizing the melodies seems to have differently affected perceptual judgements of LT compared with NC subjects, benefiting the NC group more than the LT group. This result, which differs from the data obtained with single tones, demonstrates the relevance of using melodies to test perceptual abilities, and of contrasting results obtained with both sets of stimuli. Despite preserved timbre discrimination of single tones (Samson and Zatorre, 1994), the multidimensional results obtained in this melodic condition indicate a relatively mild disturbance in higher-order perceptual organization. This finding provides additional information regarding the perceptual abilities of LT subjects and is broadly in agreement with the hypothesis of Whitfield (1985), who proposed that the function of the sensory cortex in general is to detect similarities among stimuli. To support this idea, Whitfield showed that auditory cortical lesions in cats destroy the detection of similarities between signals without necessarily affecting the detection of differences (discrimination). A similar dissociation has recently been reported in the human literature by the demonstration that frequency direction judgements are impaired in patients with lesions of the right Heschl's gyrus, whereas frequency discrimination is spared (Johnsrude *et al.*, 2000).

The major finding of this study concerns the analysis of the RT group ratings, which revealed a disturbed perceptual space in both conditions. The most distorted configuration was obtained with single tones, in which the temporal information was neglected in the first two dimensions, and tones were grouped according to their spectral content. Furthermore, the RT group's use of spectral information did not reflect a coherent underlying dimensionality, as the NC group's scaling solution did. These patients obviously tend to group tones according to spectral cues, but they seem to be able to only dissociate the more complex tones (i.e. eight-harmonic tones) from the one- and four-harmonic tones along the first dimension, and to distinguish pure tones (i.e. one-harmonic tones) from complex tones along the second dimension. These findings confirm the role of the right temporal lobe in tasks requiring timbre discrimination of spectral change (Milner, 1962; Tramo and Gazzaniga, 1989; Samson and Zatorre, 1994). Moreover, we found that patients with RT lesion can only distinguish extreme spectral features, either the pure or the most complex tones, providing some insight about the spectral information actually perceived by RT subjects.

The results from this experiment also indicate that damage involving the right temporal neocortex impaired the processing of temporal envelopes of tones characterizing musical timbre. This finding also parallels closely previous results obtained in timbre discrimination (Samson and Zatorre, 1994), in which RT patients showed a deficit in perceiving the temporal envelope of tones. In the present study, only the third dimension, which accounted for only 16% of the variance, was somewhat related to the temporal envelopes of tones, as it distinguished abrupt (1 ms) from more gradual (100 and 190 ms) rise times of complex tones. The configuration characterizing perceptual abilities of RT subjects in the single tones was significantly distorted, but nonetheless reflects an underlying structure. Indeed, the points corresponding to the tones were not randomly distributed in the spatial configuration, suggesting that dissimilarity judgements themselves were reliable, and faithfully reflect the perceptual disturbance associated with the RT damage.

The melodic condition resulted in a spatial configuration in RT subjects that was different from that for single tones, but was still disturbed by comparison with the NC data. The perceptual dimension differentiating tones with an abrupt rise time from tones with longer rise time seems to be more salient in the melodic condition than in the single tone condition. Indeed, the amount of variance that is explained by this dimension is almost one-third of the total variance in the melodic condition. As for the LT subjects, this finding indicates that the dimension which corresponds to temporal variation of tones is more accessible in a melodic phrase, in which tones with different durations and frequencies evolve within a music phrase, than in the single tone condition. Finally, the perceptual dimensions that were related to the spectral information reflect a very similar pattern of results to

the single tone condition. The melodic condition, which one might argue is a slightly more ecologically valid situation than the single tones, presumably offers more information to the listeners than the single tones. It facilitates the extraction of some temporal cues that were almost inaccessible when these patients had to judge isolated tones. The succession of tones of different durations in the melody may have enhanced the importance of time-related parameters, in accord with our hypothesis that the melodic condition would result in a more orderly perceptual representation of musical timbre than do single tones.

More generally, the data are consistent with recent proposals that anteroventral portions of temporal neocortex may be important for processing acoustic features that are characteristic of auditory 'objects' (Rauschecker and Tian, 2000). Both neurophysiological and anatomical data suggest that neurones in the anterior parabelt region of macaque monkeys may be sensitive to spectrotemporal cues such as those used here (Rauschecker *et al.*, 1997; Romanski *et al.*, 1999). Furthermore, such cues are also probably relevant for identifying features of voices, and recent functional imaging data support the contribution of areas within the superior temporal sulcus in voice perception (Belin *et al.*, 2000). These areas would probably have been damaged in the patients studied, and our finding of a greater deficit following right temporal resection would also be consistent with the greater selectivity of right superior temporal sulcus areas for voice processing observed by Belin and colleagues.

The results of the present study replicate and extend previous data suggesting an important role for the right temporal neocortex in perception of both spectral and temporal envelopes in timbre (Milner, 1962; Tramo and Gazzaniga, 1989; Samson and Zatorre, 1994), and provide additional, qualitative information about the perceptual deficits characterizing patients with right or left temporal lobe lesion. The results of this study also showed that reliable data can be obtained from subjective judgements in patients as well as in normal listeners who were all musically unsophisticated. They indicate that lesions of the left temporal lobe do affect subtle aspects of timbre perception, as measured by ratings of dissimilarity, despite these patients' preserved ability to make discrimination judgements using a traditional paradigm. Moreover, these data provide qualitative information about the perceptual abilities of RT subjects, which was not available from conventional discrimination measures. The multidimensional scaling method based on dissimilarity judgements appears to be an appropriate tool to investigate perception of multidimensional perceptual attributes such as musical timbre. This attempt was successful and indicates that such methodologies are well suited to describe dysfunctional perceptual abilities. It also allows the identification of the perceptual dimensions that remain accessible, providing therefore an additional opportunity to approach neuropsychological issues related to human perception.

Acknowledgements

References

American Standards Association. American standard acoustical terminology. S1.1–1960. New York: American Standards Association; 1960.

Balzano GJ. What are musical pitch and timbre? Music Percept 1986; 3: 297–314.

Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure M-C, et al. Lateralization of speech and auditory temporal processing. J Cogn Neurosci 1998; 10: 536–40.

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature 2000; 403: 309–12.

Berger KW. Some factors in the recognition of timbre. J Acoust Soc Am 1964; 36: 1888–91.

Carmon A, Nachshon I. Effect of unilateral brain damage on perception of temporal order. Cortex 1971; 7: 410–18.

Celesia GG, Broughton RJ, Rasmussen T, Branch C. Auditory evoked responses from the exposed human cortex. Electroencephalogr Clin Neurophysiol 1968; 24: 458–66.

De Ribaupierre F. Acoustical information processing in the auditory thalamus and cerebral cortex. In: Ehret G, Romand R, editors. The central auditory system. New York: Oxford University Press; 1997. p. 317–97.

Divenyi PL, Robinson AJ. Nonlinguistic auditory capabilities in aphasia. Brain Lang 1989; 37: 290–326.

Efron R. Temporal perception, aphasia and déjà vu. Brain 1963; 86: 403–24.

Ehrlé N, Samson S, Baulac M. Processing of rapid auditory information in epileptic patients with left temporal lobe damage. Neuropsychologia 2001; 39: 525–31.

Galaburda A, Sanides F. Cytoarchitectonic organization of the human auditory cortex. J Comp Neurol 1980; 190: 597–610.

Gordon JW, Grey JM. Perceptual effects of spectral modifications on orchestral instrumental tones. Comp Music J 1978; 2: 24–31.

Grey JM. Multidimensional perceptual scaling of musical timbres. J Acoust Soc Am 1977; 61: 1270–7.

Grey JM, Gordon JW. Perceptual effects of spectral modifications on musical timbres. J Acoust Soc Am 1978; 63: 1493–500.

Helmoltz H. Théorie physiologique de la musique fondée sur l'étude des sensations auditives. Paris: Masson; 1868.

Henke WL. An interactive dialogue language for time signal processing. Cambridge (MA): MIT Research Laboratory of Electronics; 1976.

Johnsrude IS, Penhune VB, Zatorre RJ. Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 2000; 123: 155–63.

Krimphoff J, McAdams S, Winsberg S. Charactérisation du timbre des sons complexes. II: Analyses acoustiques et quantification psychoacoustique. J Phys (Paris) 1994; 4: 625–8.

Kruskall JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 1964; 29: 1–27.

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

McAdams S, Cunible JC. Perception of timbral analogies. Philos Trans R Soc Lond B Biol Sci 1992; 336: 383–9.

McAdams S, Winsberg S, Donnadieu S, De Soete G, Krimphoff J. Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. Psychol Res 1995; 58: 177–92.

Miller JR, Carterette EC. Perceptual space for musical structures. J Acoust Soc Am 1975; 58: 711–20.

Milner B. Laterality effects in audition. In: Mountcastle VB, editor. Interhemispheric relations and cerebral dominance. Baltimore: Johns Hopkins; 1962. p. 177–95.

Pandya DN. Anatomy of the auditory cortex. [Review]. Revue Neurol Paris 1995; 151: 486–94.

Plomp R. Timbre as a multidimensional attribute of complex tones. In: Plomp R, Smoorenburg GF, editors. Frequency analysis and periodicity detection in hearing. Lieden: Sijthoff; 1970. p. 397–414.

Ramsay JO. Some statistical approaches to multidimensional scaling data. J R Stat Soc A 1982a; 145: 285–312.

Ramsay JO. MULTISCALE II manual. Montreal: McGill University, Department of Psychology; 1982b.

Rauschecker JP, Tian B. Mechanisms and streams for processing of 'what' and 'where' in auditory cortex. Proc Natl Acad Sci USA 2000; 97: 11800–6.

Rauschecker JP, Tian B, Pons T, Mishkin M. Serial and parallel processing in rhesus monkey auditory cortex. J Comp Neurol 1997; 382: 89–103.

Robin DA, Tranel D, Damasio H. Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. Brain Lang 1990; 39: 539–55.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 1999; 2: 1131–6.

Samson S, Zatorre RJ. Contribution of the right temporal lobe to musical timbre discrimination. Neuropsychologia 1994; 32: 231–40.

Samson S, Zatorre RJ, Ramsay JO. Multidimensional scaling of synthetic musical timbre: perception of spectral and temporal characteristics. Can J Exp Psychol 1997; 51: 307–15.

Samson S, Ehrlé N, Baulac M. Cerebral substrates for musical temporal processes. [Review]. Ann NY Acad Sci 2001; 930: 166–78.

Schouten JF. The perception of timbre. In: Report of the 6th International Congress on Acoustics. Tokyo: 1968. p. GP6–2.

Seashore CE, Lewis D, Saetveit JG. Revision of the Seashore measures of musical talents. Aims and progress of research, No. 65. Iowa City: University of Iowa Press; 1940.

Shepard RN. The analysis of proximities: multidimensional scaling with an unknown distance function. I. Psychometrika 1962; 27: 125–40.

Sidtis JJ, Volpe BT. Selective loss of complex-pitch or speech discrimination after unilateral lesion. Brain Lang 1988; 34: 234–45.

Swisher L, Hirsh IJ. Brain damage and the ordering of two temporally successive stimuli. Neuropsychologia 1972; 10: 137–52.

Tallal P, Newcombe F. Impairment of auditory perception and language comprehension in dysphasia. Brain Lang 1978; 5: 13–34.

Tallal P, Miller S, Fitch RH. Neurobiological basis of speech: a case for the preeminence of temporal processing. [Review]. Ann NY Acad Sci 1993; 682: 27–47.

Tramo MJ, Gazzaniga MS. Discrimination and recognition of complex tonal spectra by the cerebral hemispheres: differential lateralization of acoustic-discriminative and semantic associative functions in auditory pattern perception [abstract]. Soc Neurosci Abstr 1989; 15: 1060.

Wada J, Rasmussen T. Intracarotid injection of Sodium Amytal for the lateralization of cerebral speech dominance: experimental and clinical observations. J Neurosurg 1960; 17: 266–82.

Wechsler D. Wechsler memory scale revised manual. San Antonio (TX): Psychological Corporation; 1987.

Wessel DL. Psychoacoustics and music. Bull Comp Arts Soc 1973; 30: 1–2.

Wessel DL. Timbre space as a musical control structure. Comp Mus J 1979; 3: 45–52.

Whitfield IC. The role of auditory cortex in behavior. In: Peters A, Jones EG, editors. Cerebral cortex, Vol. 4. New York: Plenum Press; 1985. p. 329–49.

Zatorre RJ. Pitch perception of complex tones and human temporal-lobe function. J Acoust Soc Am 1988; 84: 566–72.

Zatorre RJ. Perceptual asymmetry on the dichotic fused words test and cerebral speech lateralization determined by carotid sodium amytal test. Neuropsychologia 1989; 27: 1207–19.

Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimintation in speech processing. Science 1992; 256: 846–9.