The functional anatomy of central auditory processing

Thomas E Cope,1 David M Baguley,2 Timothy D Griffiths3

INTRODUCTION
Disorders of audition stand at a crossroads between neurology, audiology, psychiatry and ENT. It is not uncommon for patients to ‘fall between the cracks’ of the referral pathway, and to become increasingly frustrated when specialties cannot decide who is responsible for them. The aim of this ‘How to understand it’ article is to demystify central auditory disorders for neurologists, and to show that they can be assessed like any other neurological symptom, based on a consideration of the anatomy and physiology of the auditory pathway. To illustrate this we describe the case history of a patient who had lost the ability to recognise and appreciate music who presented to an audiology clinician (DMB), who then sought neurology support (TEC, TDG). We elucidated the nature of the deficit in the analysis of sound patterns using psychophysics, but the deficits could have been predicted by a consideration of the anatomy and physiology of the auditory system.

CASE HISTORY
The patient was a 48-year-old, right-handed salesman who was a keen listener to popular music. He had no previous symptoms relating to environmental sound or speech perception.

He suffered two low-speed road traffic accidents in a single journey on the way to work. The first was nose-to-tail, as he failed to see the car in front stop in traffic, and the second was nose-to-nose with a parked car. After the second incident, he pulled over and called an ambulance, as he felt ‘disorientated’. The ambulance crew found him to be confused and complaining of headache and nausea, so brought him to hospital. Initial physical examination showed no injuries but his blood pressure was 230/120 mm Hg. He described no new symptoms related to his hearing after the first episode.

Four years later, he was woken from sleep by a tremendously loud noise, “like a spaceship landing”. He was disorientated, vomited and complained of headache and numbness of the right arm. He attended hospital, where it quickly became clear that he was completely deaf. Neurological examination was normal except for diplopia on left gaze, and his systolic blood pressure was 156/95 mm Hg. Over the next 12 h, his blood pressure rose to 214/76 mm Hg. Unenhanced CT of the head showed an extensive lobar haemorrhage, involving most of the right temporal lobe with some parietal lobe extension (figure 1). Digital subtraction angiography of brain was normal. He required 10 days in the high-dependency unit, primarily to control his elevated blood pressure, followed by rehabilitation on the stroke ward. He was discharged on day 28 with residual left hemianopia/neglect, and eventually made it back to work on reduced duties. Sometime later, 24-h ambulatory blood pressure monitoring was normal while taking amlodipine and ramipril. He described no new symptoms related to his hearing after the first episode.

By discharge on day 14, he was aware of noises occurring but could not characterise them. He had no comprehension of speech, and suffered extremely loud and troublesome bilateral tinnitus.
RECOVERY OF AUDITORY FUNCTION

Over the next few months, his pure-tone audiometry (a test of hearing sensitivity) gradually improved until it was in the normal range for his age. Auditory brainstem-evoked responses (figure 3) were unchanged in the 4 years after the second event, showing a normal response to left-sided clicks, but absent waves IV and V (which are generated by the superior olivary complex and inferior colliculus in upper brainstem) to right-sided clicks.

Following audiology advice, during the 4 years after the second haemorrhage, he managed to retrain his understanding of speech by reading printed books while listening to the corresponding audiobook. This coincided with a gradual improvement in speech reception threshold (the minimum volume at which a patient can understand 50% of simple consonant-vowel-consonant words). More detailed testing of speech perception 4 years after the second event found a significantly impaired ability to distinguish minimal pairs of real words (eg, bear vs pear) but excellent performance distinguishing words from non-words (eg, bus vs mus).

In contrast to speech, 4 years after the second haemorrhage he still could not recognise music. Musical deficits were confirmed then by his performance on the Montreal battery of assessment of amusia. This assesses melodic and rhythmical structure systematically with a same–different task applied to pairs of novel melodies. The patient had a borderline performance on a test of contour (the pattern of pitch rises and falls) and was below cut-off for interval (degree of pitch change), scale (keeping in the same key), metre (march or waltz), rhythm and melodic memory (table 1).

Testing of uncued environmental sound recognition showed a small number of errors (table 2).
At the time of assessment, 4 years after the second haemorrhage, he therefore had deficits in auditory perception affecting speech, music and environmental sounds in the absence of deafness: an auditory agnosia.

AUDITORY PSYCHOPHYSICS

Complex percepts like speech and music rely on the analysis and integration of a variety of auditory cues or ‘building blocks’. Deficits in the analysis of particular sensory cues can produce deficits in the perception of complex sounds that contain that cue (eg, speech or music) in the absence of deafness. We suspected he had an ‘apperceptive’ agnosia. The presence of deficits that cross domains (here affecting music, speech and environmental sounds in descending order of severity) is common in apperceptive agnosia. To tease this apart further, the patient performed a range of psychophysical tasks that test the perception of different types of sound patterns and we compared this with group norms (demographic data for control groups in table 3). This establishes a profile of performance for the perception of different sound patterns, a bit like the neuropsychological profile measured in cognitive cases. The tasks used are illustrated in cartoon form in figure 4.

We first assessed pitch analysis. The perception of the pitch of individual notes is a complex process that depends critically on the fine timing information in sound at the level of milliseconds.\(^2\) His discrimination of pitch was strikingly abnormal (<0.1st centile), which is a sufficient low-level cause for the problems in melodic perception detected by the Montreal battery of assessment of amusia using more musical test stimuli.

The time structure of sound at the level of 10s of milliseconds is a critical feature relevant to speech and music perception. He was almost completely unable to detect frequency modulation at a rate of 40 Hz that is usually heard as a roughness (unmeasurably poor performance).

Tests of timing analysis in the 100s of milliseconds range found only mild deficits. Detection of a 2-Hz frequency modulation that is usually heard as a slow vibrato was borderline (2nd centile), a high-level test of rhythm perception. In the same 100s of milliseconds range, he had normal discrimination of time intervals (54th centile), normal detection of the beat in a rhythmic sequence (17th centile) and normal metre detection (11th centile).

When asked to discriminate the spectral density of dynamic ripples (the ‘building blocks of speech’)—essentially a higher level test of timbre perception in a difficult listening environment—he performed supra-normally (>99th centile).

### Table 1: Scores on the Montreal battery of assessment of amusia

<table>
<thead>
<tr>
<th>Task</th>
<th>Scale</th>
<th>Contour</th>
<th>Interval</th>
<th>Rhythm</th>
<th>Metre</th>
<th>Memory</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>17</td>
<td>22</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>103</td>
</tr>
<tr>
<td>Cut-off score</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>20</td>
<td>22</td>
<td>140</td>
</tr>
<tr>
<td>Amusics mean Peetz et al(^1)</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>22</td>
<td>20</td>
<td>20</td>
<td>–</td>
</tr>
</tbody>
</table>

*All tasks are scored out of 30 and are 2-alternative forced-choice, such that 15/30 represents chance performance. The first four tasks have catch (easy) trials to ensure concentration, and this patient answered all of these correctly.*

### Table 2: Verbatim responses of the patient when asked to identify exemplars from a battery of sound effects

<table>
<thead>
<tr>
<th>Sound effect</th>
<th>Response quotation</th>
<th>Correct?</th>
</tr>
</thead>
<tbody>
<tr>
<td>General: Telephone</td>
<td>Bell, buzzer … telephone?</td>
<td>Yes</td>
</tr>
<tr>
<td>Door opening + closing</td>
<td>Fireworks, chainsaw … felling a tree?</td>
<td></td>
</tr>
<tr>
<td>Cash register</td>
<td>No idea</td>
<td>Yes</td>
</tr>
<tr>
<td>Crowd talking</td>
<td>Crowded room</td>
<td></td>
</tr>
<tr>
<td>Dentist’s drill</td>
<td>The edge of a jungle</td>
<td></td>
</tr>
<tr>
<td>Doorbell</td>
<td>Doorbell</td>
<td>Yes</td>
</tr>
<tr>
<td>Water pouring</td>
<td>Flushing toilet</td>
<td>Close</td>
</tr>
<tr>
<td>Pneumatic drill</td>
<td>Roadworks … jackhammer</td>
<td></td>
</tr>
<tr>
<td>Typerwriter</td>
<td>Tap dancer</td>
<td>Close</td>
</tr>
<tr>
<td>Car engine + wipers</td>
<td>Factory floor</td>
<td></td>
</tr>
<tr>
<td>Environment: Rain</td>
<td>Rain</td>
<td>Yes</td>
</tr>
<tr>
<td>Wind</td>
<td>Wind</td>
<td>Yes</td>
</tr>
<tr>
<td>Gentle waves</td>
<td>Thunderstorm</td>
<td></td>
</tr>
<tr>
<td>Thunder</td>
<td>Louder thunderstorm</td>
<td>Yes</td>
</tr>
<tr>
<td>Fog (horn)</td>
<td>No idea</td>
<td></td>
</tr>
<tr>
<td>Animals: Sheep</td>
<td>Lamb</td>
<td>Yes</td>
</tr>
<tr>
<td>Dog</td>
<td>Dog</td>
<td>Yes</td>
</tr>
<tr>
<td>Lion</td>
<td>No idea</td>
<td></td>
</tr>
<tr>
<td>Birdsong</td>
<td>Birds</td>
<td>Yes</td>
</tr>
<tr>
<td>Cat</td>
<td>Cat</td>
<td>Yes</td>
</tr>
<tr>
<td>Cockerel crowing</td>
<td>Chicken … cock to be precise</td>
<td>Yes</td>
</tr>
<tr>
<td>Cow</td>
<td>Cow</td>
<td>Yes</td>
</tr>
<tr>
<td>Gorilla</td>
<td>Horse? Heavy breathing animal … could be any animal</td>
<td>Close</td>
</tr>
<tr>
<td>Horse</td>
<td>Not any inkling whatever</td>
<td></td>
</tr>
<tr>
<td>Whale song</td>
<td>Elephant</td>
<td></td>
</tr>
</tbody>
</table>

### How to Understand It

analysis of sounds with a time structure that evolves over milliseconds or 10s of milliseconds. To understand this, we must briefly consider the anatomy of the auditory pathway (figure 5). Inner hair cells in the cochlea transduce the mechanical energy of sound into electrical impulses in the auditory nerve which, for sounds below a few kilohertz in frequency, are phase-locked (ie, nerves fire at a fixed point in the cycle of a sound wave). After ipsilateral processing in either the dorsal or the ventral cochlear nucleus—and sometimes the nucleus of the trapezoid body—impulses are projected bilaterally, but with a contralateral dominance, to the superior olivary complex. The pathway travels up through the lateral lemniscus to inferior colliculus where there is a further partial decussation. There is then a further synapse in the medial geniculate nucleus before cortical processing in primary auditory cortex on Heschl’s gyrus of the medial temporal lobe. It is important to remember that, in contrast to the visual system, there is significant signal processing at each nucleus in the pathway.

As a result of this patient’s initial haemorrhage, involving most of the right temporal lobe, we must assume that his cortical processing of auditory information is exclusively left sided. The second haemorrhage damaged the left internal colliculus, the highest level at which there is decussation of auditory tracts. This initially resulted in complete cortical deafness—as the left auditory cortex lost its meaningful input—and severe tinnitus, presumably as a result of disordered neuronal output from the internal colliculus. Over time, plasticity and perceptual retraining allowed him to make use of surviving ascending pathways to restore some useful hearing, but with permanent disruption of the synchronisation and temporal fidelity of these signals.

The cochlea represents the frequency spectrum of sounds in two ways, by the position on the basilar membrane with the largest amplitude of deflection, and by entraining neural firing to a fixed point in the cycle of a sound wave, so-called ‘phase locking’. Unlike in the visual system, where most postretinal processing occurs in the neocortex, auditory information is highly preprocessed by a series of brainstem nuclei before reaching the cortex. These nuclei are sensitive to temporal regularity, even when spectral information

Table 3 Baseline characteristics of the patient, and the means of a group of 12 normal elderly controls from whom the auditory processing norms were established, and a group of 11 patients with peripheral neuropathies from whom the timing task norms were established

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>WTAR</th>
<th>MMSE</th>
<th>ACE-R</th>
<th>Att</th>
<th>Mem</th>
<th>Flu</th>
<th>Lan</th>
<th>Vis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>56</td>
<td>42</td>
<td>30</td>
<td>99</td>
<td>18</td>
<td>25</td>
<td>14</td>
<td>26</td>
<td>16</td>
</tr>
<tr>
<td>Auditory processing controls</td>
<td>73</td>
<td>42</td>
<td>29</td>
<td>95</td>
<td>18</td>
<td>24</td>
<td>12</td>
<td>26</td>
<td>16</td>
</tr>
<tr>
<td>Timing task controls</td>
<td>60</td>
<td>42</td>
<td>29</td>
<td>94</td>
<td>18</td>
<td>24</td>
<td>12</td>
<td>25</td>
<td>16</td>
</tr>
</tbody>
</table>

ACE-R, Addenbrooke’s Cognitive Examination; Revised edition, followed by subscores in attention, memory, fluency, language, visuospatial; MMSE, mini mental state examination; WTAR, Wechsler test of adult reading.

Figure 5 Schematic diagram of the auditory pathway from cochlea to primary auditory cortex. Adapted with permission from Hill. NTB, nucleus of the trapezoid body.
is absent. Phase locking becomes progressively less precise as one ascends the auditory pathway, as this information is ‘distilled’ for further processing, freeing the neocortex from the necessity of encoding ‘pitch’ through regularity and allowing for the oscillatory activity common to other perceptual domains. Lesions to the brainstem, as in our patient, can therefore cause impairments in pitch discrimination. Because it is the first point of convergence of the auditory pathway, lesions of the brainstem also commonly lead to problems with integrating the inputs of both ears for spatial sound analysis.

In our patient, his profound deficit for detecting acoustic features over timescales of milliseconds or 10s of milliseconds correctly localises the lesion to the auditory nerve or brainstem nuclei. Deficits in temporal processing are unusual in brainstem lesions because of the incomplete decussation of the auditory pathway, which is about 70% crossed, meaning that lesions generally need to be bilateral to produce deficits. Bilateral brainstem lesions compatible with life are unusual but do occur in multiple sclerosis, in which there are well-described deficits in temporal sound analysis.

There is a general principle in the auditory system that the analysis of increasingly complex acoustic features over progressively longer time windows occurs at successive stages of the auditory pathway, and involves the recruitment of increasingly distributed networks. Here we show abnormal analyses at timescales shorter than 10s of milliseconds due to a brainstem lesion. Difficulties with spectral flux, or the analysis of syllabic transitions over 10s of milliseconds localises to primary and then associated auditory cortices, as the time window lengthens, with a general trend for longer time windows to be processed in the non-dominant hemisphere. Timing at the level of 100s of milliseconds and beyond recruits a distributed network involving the cerebellum, the supplementary motor area and the basal ganglia.

**SPEECH AND MUSIC PERCEPTION**

Music is a fragile percept, and is less robust to degraded spectral structure than is speech perception. Amusia is therefore a relatively non-specific complaint in isolation, as it can result from deficits in a variety of auditory perceptual domains. This is illustrated by our patient, who performed well at tasks of timing at the level of 100s of milliseconds, the building blocks of rhythm, beat and metre, when time intervals were marked by tone pips in an otherwise silent background, but could not perform similar tasks in the Montreal battery of assessment of amusia when they were presented in a musical context.

By contrast, speech perception is a relatively redundant system. Humans can recognise speech when the auditory input is markedly degraded, especially after a period of training, and are aided by expectations, context and cross-modal cues such as lip-reading. This is illustrated by noise-vocoded speech, from which all temporal fine structure and the majority of spectral cues are removed, and which our patient could report to a normal level of accuracy.

When the difficulty with identifying and understanding music and speech is, as here, due to abnormal analysis of the structure of sounds, this represents an apperceptive auditory agnosia. In some other cases, the spectrotemporal analysis of sound is intact but there is abnormal association of the percept with meaning, which represents an associative auditory agnosia.

**CONCLUSION**

In our experience, some neurologists dismiss auditory processing as being of ‘niche’ interest, and it is true that patients such as this are uncommon. Nonetheless, difficulties with auditory processing occur in a wide variety of ‘core’ neurological conditions. Multiple
sclerosis is a common cause of brainstem lesions, and the efferent auditory pathway can be abnormal even with a normal MRI.\textsuperscript{13} Patients with aphasias resulting from stroke\textsuperscript{12} or neurodegeneration\textsuperscript{33–35} have significant difficulties with the processing of non-verbal sounds. Abnormal auditory sensations can occur in epilepsy, either as aura\textsuperscript{16} or after seizures,\textsuperscript{37} leading to a diagnosis of `focal epilepsy with auditory features',\textsuperscript{38} which can be sporadic or related to mutation of the LGI1 `epitempin' gene.\textsuperscript{39} Equally, emotionally valent sounds, and especially music, can precipitate seizures as part of `reflex epilepsies'.\textsuperscript{40,41} In all cases, clinicians should attempt to link structure and function.

**Contributors** DMB made the initial clinical audiological assessment and oversaw audiological recovery. He sought a neurological opinion from TEC and TDG, who assessed the patient together as an NHS service. TEC and TDG designed the psychophysical assessments and TEC administered them. All three authors prepared the manuscript.

**Funding** The assessments described were performed as part of an NHS service so this work was not directly funded, but TEC is a NIHR Academic Clinical Fellow, and TDG is a Wellcome Trust Senior Clinical Fellow (WT091681MA).

**Competing interests** None declared.

**Provenance and peer review** Commissioned; externally peer reviewed. This paper was reviewed by Chris Butler, Oxford, UK.

**Open Access** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

**Further reading**


**REFERENCES**


