How do our brains analyze temporal structure in sound?

Robert Zatorre

How does the human brain process the temporal structure of a musical sound? A PET imaging study identifies cortical regions involved in pitch computation and melodic pattern recognition.

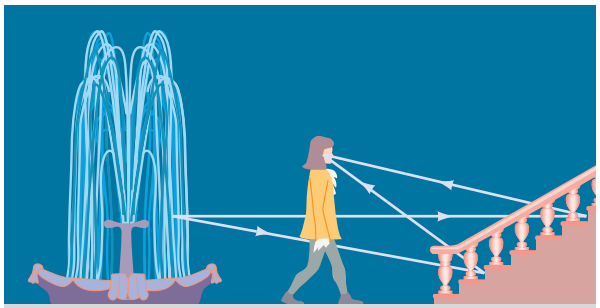
As a consequence of its physical nature, sound unfolds over time. This presents unique problems for the nervous system, which is faced with the task of extracting and encoding information contained in vibrating air molecules over short time periods, as well as integrating successive sonic events over longer time frames. The auditory system is therefore highly specialized for processing temporal events, which constitute the lowest common denominator of everything from the bark of a dog to a Bach cantata. The brain imaging study by Griffiths and colleagues in this issue of *Nature Neuroscience* (pp 422–427) examines how the human brain may process the temporal structure contained in a musical sound. They identify different cortical regions involved in the processing of short-term (pitch computation) and longer-term (melodic pattern) temporal information.

Robert Zatorre is at the Department of Neuropsychology, Montreal Neurological Hospital, 3901 University, Montreal PQ H3A 2B4 Canada
email: md37@musica.mcgill.ca

These investigators took advantage of a phenomenon described some three hundred years ago by Huygens[1], who noted that the periodic reflections of the noise made by a fountain from the stone steps of a staircase resulted in an audible pitch (Fig. 1). The same phenomenon can be reproduced and studied in the laboratory by taking a sample of random noise and passing it through a cascade of delay-and-add networks. That is, the noise is displaced by a brief time delay and added to itself; then the output of that procedure is again added to itself with the same delay, and so on for some number of iterations. This process results in an audible pitch corresponding to the reciprocal of the time delay constant used (ref. 1 and Fourcin, A., *Fifth Intl. Congress Acoust.*, B42, 1965). This pitch sensation is of particular interest because it is perceived despite the absence of any auditory spectral cues that could activate a particular frequency representation in the cochlea, in contrast to a typical complex tone. Instead, the perception of pitch must result from the central nervous system's capacity to encode temporal regularity in the neural firing pattern.

...seeking areas whose activity correlated... with increasing salience of the melody.

Since the time of Helmholtz, it has been controversial whether temporal or spectral information is the basis for the sensation of pitch. This duality arises because temporal regularity in a signal, or periodicity, gives rise both to temporally regular neur-



Bob Crimi

Fig. 1. Illustration of the phenomenon described by Christian Huygens in 1693. He noted that the noise produced by a fountain at the château of Chantilly de la Cour was reflected by a stone staircase in such a way that it produced a musical tone. He correctly deduced that this was due to the successively longer time intervals taken for the reflections from each step to reach the listener's ear.

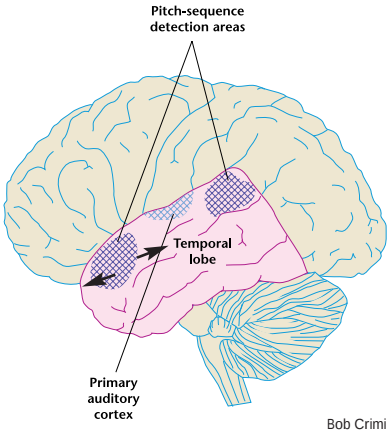
al firing patterns and to spatially distinct neural activity for different frequencies. In other words, frequency of vibration is represented by the nervous system as both a time code and a place code. Both phenomena arise in the inner ear; the time code is made possible because auditory fibers originating in the cochlea fire action potentials that are phase-locked to the stimulus, whereas the place code, known as tonotopy, arises because frequency preference is mapped along the length of the basilar membrane in the cochlea. Because of the orderly projection of fibers through the auditory pathway, this tonotopic map is maintained in the primary auditory cortex and adjacent areas[2]. It has therefore traditionally been difficult to determine whether a particular effect may arise from temporal or spectral aspects of pitch processing. The stimulus used by Griffiths and colleagues is particularly well suited to studying temporal processing in isolation, because it results in an essentially flat spectrum of cochlear activation (as shown by the simulations illustrated in their Fig. 1), with no peak at the frequency corresponding to the perceived pitch. Hence, the pitch that is perceived as a result, and any associated neural activity that is correlated with pitch perception, must result from temporal processing.

How and where does the brain extract pitch information from temporal cues? In their first experiment, Griffiths and colleagues used positron emission tomography (PET) to analyze changes in cerebral blood flow (CBF, which reflects neural activity) as a function of increasing number of iterations of the noise-delay stimulus. Each successive iteration of the stimulus sequence produces a stronger sense of pitch. This approach, known as a parametric design, is a powerful one that is becoming increasingly common in the brain imaging community. By looking for CBF changes that covary with an input variable (in this instance, number of iterations), one avoids the problems associated with subtraction analysis, in which it is necessary to assume that two experimental conditions differ only by a discrete number of components. In a parametric design, by contrast, all that is required is that the CBF changes be related in a roughly linear fashion to the input variable. Moreover, although techniques such as PET and fMRI suffer from relatively poor temporal resolution, it is nonetheless possible to use them to study temporal processing at a much finer time scale by varying the stimulus dimension of interest appropriately. That is, rather than measuring the millisecond-level timing of responses, the researcher measures the differences in response to stimuli that vary on that time scale.

As the number of iterations was increased, the authors observed a nearly linear increase in CBF bilaterally in the superior temporal gyrus, near the primary auditory cortex of Heschl's gyrus (see Fig. 2), as determined from anatomical probability maps[3]. These results directly implicate a region within or adjacent to the primary auditory cortex in processing temporal fine structure, and indicate that temporal integration probably occurs at or before the primary auditory cortex, perhaps in the inferior colliculus or medial geniculate nucleus of the thalamus. One question that arises from these data is the extent to which the CBF pattern directly correlates with perceived pitch salience, as opposed to the physical structure of the stimulus. It would be of interest, for example, to investigate how perceived pitch sensation (as rated by the subjects) relates to the observed pattern of cerebral activity. It will also be important to compare more directly the cortical mapping of temporally based pitch and of spectrally determined pitch to see to what extent they correspond.

In other functional imaging studies, specific activation patterns seem to be related to the processing of stimuli containing spectral modulation but whose temporal structure is constant (Thivard, L. *et al.*, *NeuroImage*, 7, 373, 1998). These data complement those of Griffiths and colleagues and seem to point to distinct, functionally specialized cortical fields that are concerned with spectral or temporal aspects of stimulus processing. Such findings are also in keeping with neurophysiological data showing sensitivity of neurons in the superior temporal gyrus of the macaque to stimuli with varying spectral energy peaks[4]. These peaks, or formants, are interesting because they are relevant for the perception of speech, which depends on both spectral and temporal cues. What remains to be done, however, is to specify which areas in particular are involved, what their boundaries are, and most importantly what their precise functional contribution to different types of auditory processing may be.

In their second experiment, Griffiths and colleagues extended their parametric approach to examine the CBF pattern associated with longer-term temporal integration processes. They used the same iterated noise stimuli described above, but this time the pitches formed a brief structured melodic pattern (whereas in the prior condition they only formed a relatively monotonous tonal pattern). This manipulation allowed the authors to investigate the emer-



Bob Crimi

Fig. 2. Lateral view of the human brain illustrating areas of cortex activated by delay-and-add noise. The primary auditory cortex is within the Sylvian fissure, on the internal surface of the temporal lobe. The pitch-sequence detection areas described by Griffiths and colleagues are on the external surface of the temporal lobe. These areas are activated by delay-and-add 'notes' arranged into musical phrases.

gence of longer-term structure in the stimulus by seeking areas that showed an interaction between number of iterations and the presence or absence of a melodic pattern. In other words, they were seeking areas whose activity was correlated not with the perception of pitch *per se*, but rather with the increasing salience of the melody (as its component pitches were strengthened by increasing the number of iterations). Four such areas were identified that increased their response as a function of iteration much more for tonal melodic patterns than for non-melodic pitch patterns. These regions were located symmetrically in the two hemispheres: two in the superior temporal gyrus posterior to Heschl's gyrus and two in the superior temporal sulcus/middle temporal gyrus (Fig. 2). These four regions were spatially distinct from the peri-primary areas identified in the previous analysis, suggesting that the fine temporal processing related to pitch extraction involves different regions and mechanisms than the longer-term time processing that would be involved in the detection of melodic patterns.

A number of previous studies that have examined pitch processing either by functional imaging[5,6] or by studying brain-damaged patients[7,8] have suggested that tonal processing is lateralized to the right temporal cortex. This conclusion seems to be at odds with the bilateral activation observed by Griffiths and colleagues. However, pitch processing is clearly very complex, and the various studies may have been examining different aspects of this phenomenon. For instance, the right-side bias that has emerged from some earlier studies may reflect a process based largely on spectral cues, whereas the fine-grained temporal analysis of periodicity in the present study may be a more bilaterally distributed function. Furthermore, the lesion data indicate that whereas relatively mild deficits occur after unilateral cortical damage, much more devastating effects are observed with bilateral lesions[9], hence indicating that hemispheric differences are not absolute.

Nevertheless, the differences raise a general question that arises with any functional imaging experiment. Functional imaging, as exemplified by this carefully designed and executed study, can provide crucially important information about human perceptual and cognitive functions. Nonetheless, it can only identify neural correlates of these functions, and it cannot determine the degree to which a given site of activity provides essential computations. In the long run, it will be important to combine functional imaging data with the results of lesion studies if we are to gain a clearer understanding of the complexities of the human auditory system.

1. Bilsen, F. & Ritsma, R. *Acoustica* 22, 63–73 (1969/70).
2. Merzenich, M. M. & Brugge, J. F. *Brain Res.* 60, 315–333 (1973).
3. Penhune, V. B., Zatorre, R. J., MacDonald, J. D. & Evans, A. C. *Cereb. Cortex* 6, 661–672 (1996).
4. Rauschecker, J. P., Tian, B. & Hauser, M. *Science* 268, 111–114 (1995).
5. Zatorre, R. J., Evans, A. C., Meyer, E. & Gjedde, A. *Science* 256, 846–849 (1992).
6. Zatorre, R. J., Evans, A. C. & Meyer, E. *J. Neurosci.* 14, 1908–1919 (1994).
7. Zatorre, R. J. *J. Acoust. Soc. Am.* 84, 566–572 (1988).
8. Zatorre, R. J. & Samson, S. *Brain* 114, 2403–2417 (1991).
9. Peretz, I. *et al. Brain* 117, 1283–1301 (1994).

© 1998 Nature America Inc. • http://neurosci.nature.com