

Modularity of music processing

Isabelle Peretz¹ & Max Coltheart²

The music faculty is not a monolithic entity that a person either has or does not. Rather, it comprises a set of neurally isolable processing components, each having the potential to be specialized for music. Here we propose a functional architecture for music processing that captures the typical properties of modular organization. The model rests essentially on the analysis of music-related deficits in neurologically impaired individuals, but provides useful guidelines for exploring the music faculty in normal people, using methods such as neuroimaging.

Musical ability has traditionally been studied as the product of a general-purpose cognitive architecture^{1,2}, but a growing number of studies are based on the premise that music is a cognitively unique and evolutionary distinct faculty. Musical abilities are now studied as part of a distinct mental module with its own procedures and knowledge bases that are associated with dedicated and separate neural substrates³. Thus, research concerning musical ability now tends to adhere, more or less explicitly, to the concept of modularity of cognitive functions as formulated by Fodor^{4,5}.

After briefly describing what is currently meant by modularity, we will illustrate how modularity shapes current thinking about how the mind processes music, relying particularly on evidence from individuals with abnormalities of musical ability.

Modularity

According to Fodor^{4,5}, mental modules have the following characteristic properties: rapidity of operation, automaticity, domain-specificity, informational encapsulation, neural specificity and innateness. Fodor does not insist that any one of these properties is absolutely necessary for the ascription of the term 'modular'. For example, a system can be modular even if not innate; that is why there is no difficulty in describing the reading system as a module, even though reading is clearly not an innate ability⁵. So each of these properties might best be described as a typical, rather than necessary or sufficient, feature of a modular system.

Fodor does, however, consider one property to be more important than the others: information encapsulation. By this he means that the information processing within a mental module is immune from influence by the 'central system'—a large and slowly operating encyclopedic-knowledge system involved in high-level cognitive operations, such as problem solving or belief evaluation. Our view⁶ is that domain-specificity is an equally important property: it would be very odd to describe some system as a module if its operation were not specific to some restricted domain of input or output. Hence we consider domain-specificity to be not only an essential

but *necessary* property for a processing system to be considered modular.

It is also important to realize that a module can be composed of smaller processing subsystems that can themselves be referred to as modules. For example, the language module contains component lexical and phonetic-processor modules.

To claim that there is a music-processing module is to claim that there is a mental information processing system whose operation is specific to the processing of music. That system may contain smaller modules whose processing domains may also be restricted to particular aspects of music. The possibility that such a cognitive architecture for music processing exists has been entertained for more than a decade^{4,7–9}.

Even if the human mind does contain a music module, it is conceivable that this module could lack the property of neural specificity, or neuroanatomical separability. For example, the neural substrate for music processing might overlap with that used for processing other complex patterns, such as speech sounds. In this case, brain damage would never affect musical abilities while sparing all other aspects of cognition (particularly auditory processing outside the domain of music). If, on the other hand, the putative music module does possess the property of neural specificity, then we should expect to find people in whom brain damage has selectively affected musical abilities. Many such people have been found.

A module for music processing

Support for the existence of a music-processing module can be found in reports of selective impairments in music recognition abilities after brain damage (Table 1). Such patients can no longer recognize melodies (presented without words) that were highly familiar to them before the onset of their brain damage. In contrast, they are normal at recognizing spoken lyrics (and spoken words in general), familiar voices and other environmental sounds (such as animal cries, traffic noises and human vocal sounds). This condition is called 'acquired amusia'^{10–15}. Similarly, in 'congenital amusia'¹⁶, individuals suffer from lifelong difficulties with music but can recognize the lyrics of familiar songs even though they are unable to recognize the tune that usually accompanies them¹⁷.

Most people are experts at recognizing spoken words, but amateurs at recognizing music. It might therefore be argued that there is no spe-

¹University of Montreal, Box 6128, Succ. Centre-ville, Montreal, Quebec H3C 3J7, Canada. ²Macquarie Center for Cognitive Science, Macquarie University, Sydney NSW 2109, Australia. Correspondence should be addressed to I.P. (Isabelle.Peretz@umontreal.ca).

Table 1 Case reports of selective impairment and selective sparing in the auditory recognition of words, tunes and other meaningful sounds

Reports	Auditory domains		
	Tunes	Words	Other familiar sounds
C.N. and G.L. ¹⁰	–	+	+ (+ voices)
I.R. ¹¹	–	+	+
H.V. ¹²	–	+	+
H.J. ¹³			
1 case ¹⁴			
K.B. ¹⁵			
11 cases of congenital amusia ¹⁷	–	+	+ (+ voices)
1 case ¹⁸	+	–	+
1 case, during recovery ¹⁹	+	–	+
N.S. ²⁰	+	–	+
G.L. ²¹	+	–	–
1 case ²²	+	–	–
1 case ²³	+	–	–

+ normal; – impaired.

cific module for recognizing music, but just a general auditory recognition module; when that is damaged, amateur abilities such as music recognition suffer more than expert abilities such as speech recognition. This account predicts that one will not find people in whom brain damage has impaired the ability to recognize spoken words while sparing the ability to recognize music. But such cases do exist: non-musicians may lose their ability to recognize spoken words yet remain able to recognize music^{18–20} (Table 1). Similarly, brain-damaged patients who are afflicted with verbal agnosia (or word deafness), and hence have lost their ability to recognize spoken words, can be normal at recognizing nonverbal sounds, including music^{21–23} (Table 1). The existence of such cases of selective impairment and sparing of musical abilities is incompatible with the claim that there is a single processing system responsible for the recognition of speech, music and environmental sounds. Rather, the evidence points to the existence of at least two distinct processing modules: one for music and one for speech.

Modular architecture of music processing

The study of neurological deficits has revealed far more about music processing than merely that there is a mental module specific to the processing of music. The model in Fig. 1 shows the functional architecture of music processing that has been derived from case studies of specific music impairments in brain-damaged patients (see ref. 24 for earlier versions of this model). In this model, a neurological anomaly could either damage a processing component (box) or interfere with the flow of information (arrow) between components.

Two modular aspects of the model deserve comment. First, the individuation of each box or arrow in the model arises from the study of its breakdown pattern in a brain-damaged patient. This fact confers upon the individuated component the modular property of neuroanatomical separability. Second, the model proposes various music-processing modules, each of which is concerned with a particular information-processing operation that contributes to the overall system.

An example of a distinct music-specific component inside the music module is the system concerned with tonal encoding of pitch (Fig. 1).

Central to pitch organization is the perception of pitch along musical scales. A musical scale refers to the use of a small subset of pitches (usually seven) in a given musical piece. Scale tones are not equivalent and are organized around a central tone, called the tonic. Usually, a musical piece starts and ends on the tonic. The other scale tones are arranged in a hierarchy of importance or stability, with the fifth scale tone and the third scale tone being most closely related to the tonic. The remaining scale tones are less related to the tonic, and the non-scale tones are the least related; the latter often sound like ‘foreign’ tones. This tonal hierarchical organization of pitch facilitates perception, memory and performance of music by creating expectancies²⁵.

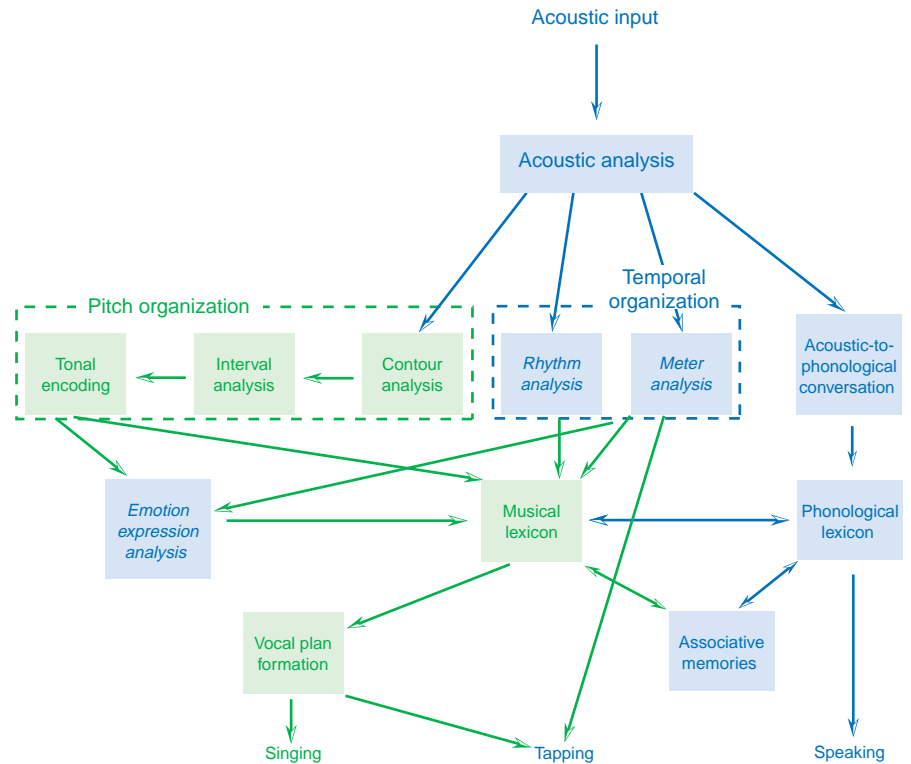
There is substantial empirical evidence that listeners use this tonal knowledge in music perception automatically^{26,27}. Tonal organization of pitch applies to most types of music, but it does not occur in processing other sound patterns, such as speech. Although the commonly used scales differ somewhat from culture to culture, most musical scales use pitches of unequal spacing that are organized around 5–7 focal pitches²⁸ and afford the building of pitch hierarchies²⁹. The tonal encoding module seems to exploit musical predispositions, as infants show enhanced processing for scales with unequal pitch steps³⁰ (see accompanying review³¹ in this issue). Tonal encoding can be selectively impaired by brain damage; for example, some patients are no longer able to judge melodic closure properly and suffer from a severe reduction in pitch memory³². In a recent functional neuroimaging study, Janata and collaborators³³ point to the rostromedial prefrontal cortex as a likely brain substrate for the ‘tonal encoding’ module.

Unlike the tonal encoding module, other component music-processing modules might not be restricted to just music (blue boxes in Fig. 1). For example, the ‘contour analysis’ component, which abstracts the pitch trajectories (in terms of pitch direction between adjacent tones without regard to the precise pitch intervals), could conceivably be involved in processing speech intonation as well as music (see accompanying review³⁴ in this issue).

The model in Fig. 1 takes as input any acoustic stimulus that can be attributed to a unique source. This implies that auditory segregation of sound mixtures into distinct sound sources first occurs in an acoustic analysis module whose domain is all auditory stimuli, not just music. The output of this early acoustic analysis might be, for example, a representation of the song “Happy Birthday.” In that case, the lyric component of the song is assumed to be processed in parallel in the language processing system (right of the figure). We suppose that activation of the music or the language processing modules is determined by the aspect of the input to which a module is tuned³⁵. That is, there is no ‘gatekeeper’ that decides which part of the auditory pattern should be sent to the musical modules and which part should be sent to the language system. All the information contained in the song line, for example, would be sent to all modules. Only the modules that are specialized for the extraction of such information will respond—just as the retina does not respond when a sound wave passes through it, nor the cochlea when light shines upon it.

The musical input modules are organized in two parallel and largely independent subsystems whose functions are to specify, respectively, the pitch content (the melodic contour and the tonal functions of the successive pitch intervals) and the temporal content, by representing the metric organization as well as the rhythmic structure of the successive durations. The ‘rhythm analysis’ component deals with the segmentation of the ongoing sequence into temporal groups on the basis of durational values without regard to periodicity; the ‘meter analysis’ component extracts an underlying temporal regularity or beat, corresponding to periodic alternation between strong and weak beats. The strong beats generally correspond to the spontaneous tapping of the

Figure 1 A modular model of music processing. Each box represents a processing component, and arrows represent pathways of information flow or communication between processing components. A neurological anomaly may either damage a processing component (box) or interfere with the flow of information between two boxes. All components whose domains appear to be specific to music are in green; others are in blue. There are three neurally individuated components in italics—rhythm analysis, meter analysis and emotion expression analysis—whose specificity to music is currently unknown. They are represented here in blue, but future work may provide evidence for representing them in green.



Ivelisse Robles

foot (thus a direct connection to tapping in Fig. 1; see accompanying review³⁶ in this issue). Both the melodic and temporal pathways send their respective outputs to either the ‘musical lexicon’ or the ‘emotion expression analysis’ component. The musical lexicon is a representational system that contains all the representations of the specific musical phrases to which one has been exposed during one’s lifetime. The same system also keeps a record of any new incoming musical input. Accordingly, successful recognition of a familiar tune depends on a selection procedure that takes place in the musical lexicon. The output of the musical lexicon can feed two different components, depending on task requirements. If the goal is to sing a song like “Happy Birthday,” the corresponding melody, represented in the musical lexicon, will be paired with its associated lyrics that are stored in the phonological lexicon and will be tightly integrated and planned in a way that is suitable for vocal production. If the task requires retrieving nonmusical information about a musical selection, such as naming the tune or retrieving a related experience from memory, the associated knowledge stored in the ‘associative memories’ component will be invoked.

In parallel with memory processes, but independently, the perceptual modules will feed their outputs into an emotion expression analysis component, allowing the listener to recognize and experience the emotion expressed by the music³⁷. This emotional pathway also contributes to recognition via the musical lexicon. Emotion expression analysis is a pivotal processing component because music has the power to elicit strong emotional responses. It takes as input emotion-specific musical features, such as mode (*e.g.* major or minor) and tempo (*e.g.* slow or fast) as computed by the melodic and temporal pathways, respectively. What is currently unclear is to what extent this emotion expression analysis component is specific to music as opposed to being involved in more general kinds of emotional processing. A patient who could recognize pieces of music but could not respond emotionally to them, while being able to respond emotionally to other media, would be informative here.

In sum, we propose a modular functional architecture for music processing that comprises several component modules. Our model (Fig. 1) also describes the pathways of information flow among these component modules. The characterization of each box and arrow represented in the model has been provided by the detailed study of brain-damaged patients with selective impairments or preservations of particular musical abilities (for review, see ref. 24). The inclusion of

three new output modules again stems from the study of neurological patients: singing performance in aphasic patients³⁸ and tapping abilities in adults suffering from congenital amusia³⁹. Thus, our proposed modular architecture for processing music provides a plausible framework for further investigating the neural mechanisms of music processing.

ACKNOWLEDGMENTS

Based on research supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research to I.P. We thank C. Palmer and T. Griffiths for insightful comments made on an earlier draft.

Received 18 March; accepted 21 April 2003
Published online 25 June 2003; doi:10.1038/nn1083

- Handel, S. *Listening: an Introduction to the Perception of Auditory Events* (MIT press, Cambridge, Massachusetts, 1989).
- Bregman, A. *Auditory Scene Analysis. The Perceptual Organization of Sound.* (MIT press, London, 1990).
- Zatorre, R. & Peretz, I. (eds.) *The Biological Foundations of Music* (NY Acad. Sci., New York, 2001).
- Fodor, J. *The Modularity of Mind* (MIT press, Cambridge, Massachusetts, 1983).
- Fodor, J. *The Mind Doesn't Work That Way* (MIT press, Cambridge, Massachusetts, 2001).
- Coltheart, M. Modularity and cognition. *Trends Cogn. Sci.* **3**, 115–120 (1999).
- Gardner, H. Musical intelligence. in *Frames of Mind* (ed. Gardner, H.) 31–58 (Basic Books, New York, 1983).
- Jackendoff, R. *Consciousness and the Computational Mind* (MIT Press, Cambridge, Massachusetts, 1987).
- Peretz, I. & Morais, J. Music and modularity. *Contemporary Music Rev.* **4**, 277–291 (1989).
- Peretz, I. *et al.* Functional dissociations following bilateral lesions of auditory cortex. *Brain* **117**, 1283–1302 (1994).
- Peretz, I., Belleville, S. & Fontaine, S. Dissociations entre musique et langage après atteinte cérébrale: un nouveau cas d'amusie sans aphasie. *Can. J. Exp. Psychol.* **51**, 354–368 (1997).
- Griffiths, T.D. *et al.* Spatial and temporal auditory processing deficits following right hemisphere infarction: a psychophysical study. *Brain* **120**, 785–794 (1997).
- Wilson, S.J. & Pressing, J. Neuropsychological assessment and the modeling of

- musical deficits. in *Music Medicine and Music Therapy: Expanding Horizons* (eds. Pratt, R.R. & Erdonmez Grocke, D.) 47–74 (Univ. Melbourne Press, Melbourne, 1999).
14. Piccirilli, M., Sciarra, T. & Luzzi, S. Modularity of music: evidence from a case of pure amusia. *J. Neurol. Neurosurg. Psychiatry* **69**, 541–545 (2000).
 15. Steinke, W.R., Cuddy, L.L. & Jakobson, L.S. Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. *Cognit. Neuropsychol.* **18**, 411–437 (2001).
 16. Peretz, I. *et al.* Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* **33**, 185–191 (2002).
 17. Ayotte, J., Peretz, I. & Hyde, K. Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain* **125**, 238–251 (2002).
 18. Laignel-Lavastine, M. & Alajouanine, T. Un cas d'agnosie auditive. *Société de Neurologie* **37**, 194–198 (1921).
 19. Godefroy, O. *et al.* Psychoacoustical deficits related to bilateral subcortical hemorrhages. A case with apperceptive auditory agnosia. *Cortex* **31**, 149–159 (1995).
 20. Mendez, M. Generalized auditory agnosia with spared music recognition in a left-hander. Analysis of a case with a right temporal stroke. *Cortex* **37**, 139–150 (2001).
 21. Metz-Lutz, M.N. & Dahl, E. Analysis of word comprehension in a case of pure word deafness. *Brain Lang.* **23**, 13–25 (1984).
 22. Takahashi, N. *et al.* Pure word deafness due to left hemisphere damage. *Cortex* **28**, 295–303 (1992).
 23. Yaqub, B.A., Gascon, G.G., Al-Nosha, M. & Whitaker, H. Pure word deafness (acquired verbal auditory agnosia) in an Arabic speaking patient. *Brain* **111**, 457–466 (1988).
 24. Peretz, I. Music perception and recognition. in *The Handbook of Cognitive Neuropsychology* (ed. Rapp, B.) 519–540 (Psychology Press, Hove, UK, 2001).
 25. Krumhansl, C.L. *Cognitive Foundations of Musical Pitch* (Oxford Univ. Press, New York, 1990).
 26. Shepard, R. & Jordan, D. Auditory illusions demonstrating that tones are assimilated to an internalized musical scale. *Science* **226** 1333–1334 (1984).
 27. Tillmann, B., Bharucha, J.J. & Bigand, E. Implicit learning of tonality: a self-organizing approach. *Psychol. Rev.* **107**, 885–913 (2000).
 28. Dowling, W.J. Melodic information processing and its development. in *The Psychology of Music* (ed. Deutsch, D.) 413–430 (Academic, New York, 1982).
 29. Balzano, G. The pitch set as a level of description for studying musical pitch perception. in *Music, Mind and Brain* (ed. Clynes, M.) 321–351 (Plenum, New York, 1982).
 30. Trehub, S.E., Schellenberg, E.G. & Kamenetsky, S.B. Infants' and adults' perception of scale structure. *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 965–975 (1999).
 31. Trehub, S.E. The developmental origins of musicality. *Nat. Neurosci.* **6**, 669–673 (2003).
 32. Peretz, I. Auditory atonalia for melodies. *Cognit. Neuropsychol.* **10**, 21–56 (1993).
 33. Janata, P. *et al.* The cortical topography of tonal structures underlying Western music. *Science* **298**, 2167–2170 (2002).
 34. Patel, A. Language, music and the brain. *Nat. Neurosci.* **6**, 674–681(2003).
 35. Coltheart, M. Assumptions and methods in cognitive neuropsychology. in *The Handbook of Cognitive Neuropsychology* (ed. Rapp, B.) 3–21 (Psychology Press, Hove, UK, 2001).
 36. Janata, P. & Grafton, S. Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat. Neurosci.* **6**, 682–687 (2003).
 37. Peretz, I. Listen to the brain: the biological perspective on musical emotions. in *Music and Emotion: Theory and Research* (eds. Juslin, P. & Sloboda, J.) 105–134 (Oxford Univ. Press, 2001).
 38. Hébert, S., Racette, A., Gagnon, L. & Peretz, I. Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain* (in press).
 39. Dalla Bella, S. & Peretz, I. Congenital amusia interferes with the ability to synchronize with music. *Ann. NY Acad. Sci.* (in press).