Understanding Bach, 2, 9-18

Bach in Mind

IAN CROSS

Bach's music has provided opportunities, posed challenges and set problems for the cognitive sciences of music. In some respects, many aspects that appeared problematic twenty years ago are perhaps on the brink of resolution, whereas other problems are only now beginning to be recognised. In this paper I attempt to give a condensed overview of how two specific domains of the cognitive sciences of music have approached the music of J. S. Bach, those domains being computational modelling and experimental research. I shall use a central example in each domain to try to bring out the achievements and failures of the cognitive sciences in advancing our understanding of Bach's music.

In this paper, the images of Bach's music that have been used in research on the cognitive sciences of music are discussed, rather than the music itself. These images have tended to take as their premise the idea that Bach's music constitutes corpora of abstract rule-governed entities that can be explored either in terms of the abstract formal principles that might underlie them or the principles that might govern the human perception of the world in sound and music. In the first case, an image of Bach's music (usually, an image that is manifested only in western common practice period notation) is addressed as abstract, mathematical, and complex; a challenge which, by virtue of the size of the corpus, seems manageable. In the second case, the image addressed (and here it is manifested in sound) is in terms of patterns that embody and exemplify generic principles of the workings of the human mind.

The roots of cognitive science are based on the premise that it might be possible to understand the human mind in the same terms as the operation of a computer. This is not to claim that the mind is a computer or that a computer is a mind. Computational theory simply provides a tool for the exploration of the human mind, either by attempting to provide computational models that seem to embody human behaviours or by testing the predictions of formal or computational theories against real human behaviours.

Bach's music seems to constitute a 'natural' focus for the cognitive sciences, in that Bach was prolific and his music seems complex yet 'lawful'. He produced much music that could be classified as being instances of the same kind, such as the chorales, thus enabling theories to be tested over a large population of instances. That his music seems both complex and lawful is also of interest, because much of it appears to exhibit patternings which, although complicated, conform to specifiable sets of rules, and thus may be explained in terms of formal principles which might be implemented computationally.

Modelling studies

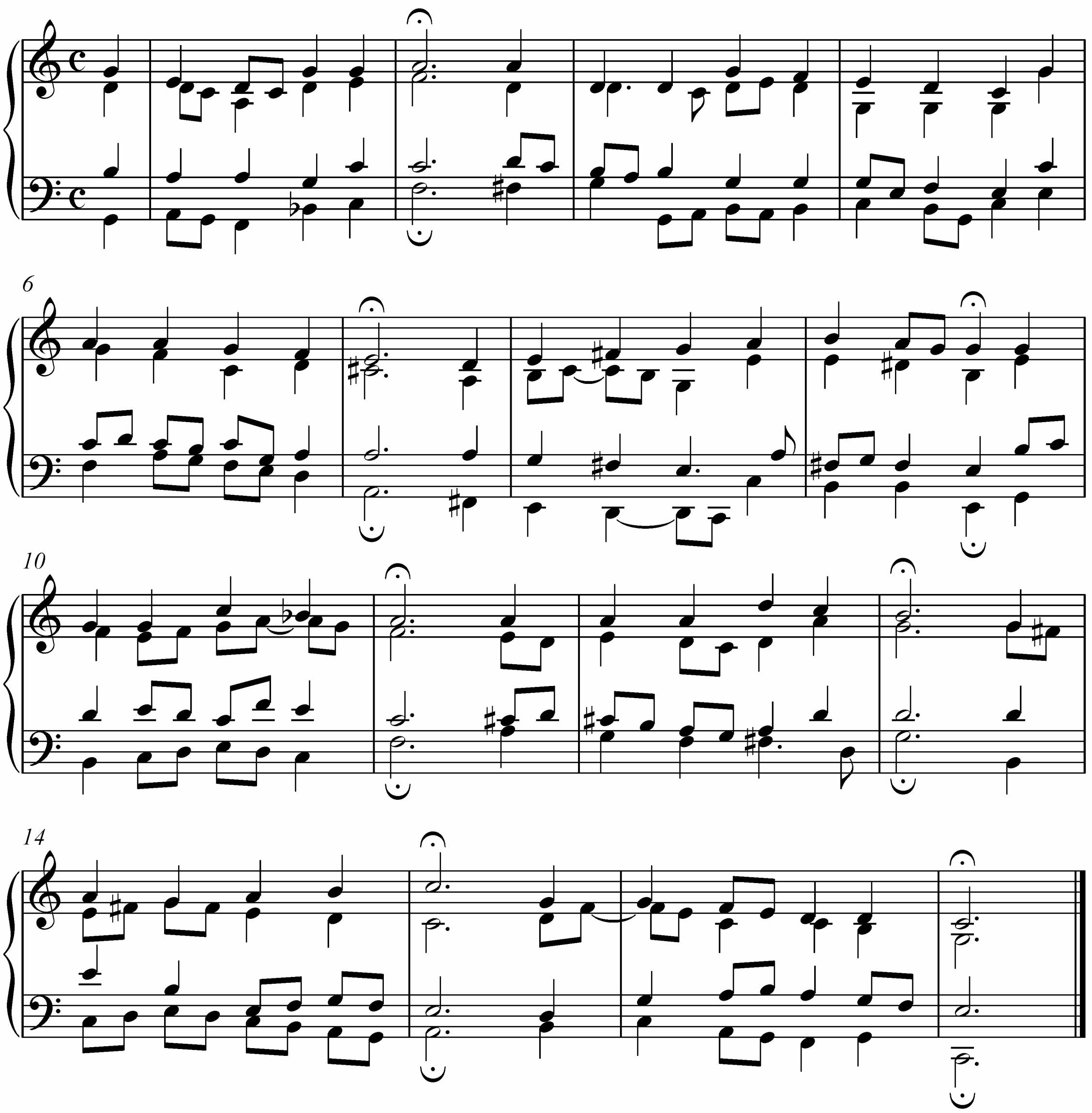
To start with, I shall consider an attempt to model the process of harmonising chorales in the style of Bach. In the 1980s, Kemal Ebcioglu produced a program that 'harmonises chorales in the style of Bach'.[1] It is an 'expert' musical system, i.e. one which incorporates a knowledge-base and addresses a tightly constrained set of problems. Its generalisability in the domain of music is limited, although its generalisability in the computational domain might be much greater. The system is symbolic, in that it explicitly encodes representations of the domain that it is attempting to model and operates according to a set of principles of formal logic.

The knowledge-base for Ebcioglu's program incorporates information about 'appropriate' harmonic and melodic structures and voice-leadings, which take the form of multiple viewpoints – one that observes the chord skeleton of the chorale, one that observes the melodic lines of the individual voices, and one that observes Schenkerian voice-leading between soprano and bass. Different viewpoints may be more or less dependent on one another for different cases. These viewpoints operate across:
i) a set of production rules (taking the following form: IF certain conditions are true of a partial solution to the problem, THEN a certain value can be assigned to the partial solution)
ii) a set of formal constraints (which narrow the range of 'legitimate' solutions)
iii) a set of heuristics (which guide the search for a solution by assigning different values to different possible solutions – in other words, the heuristic component imposes non-formal observations about the desirability of certain solutions for chorale harmonisation on the solutions generated by the production rules and constraints).

The CHORAL program works by a 'generate and test' method and (in 1992) took from half-an-hour to several hours to harmonise a chorale. This was one of the few successful ‘autonomous’ music generation programs to employ a symbolic approach, and one that made an attempt to represent attributes of music not evidently encoded directly in notation (although these are music-theoretic primitives). The program produces stylistically homogeneous and recognisably Bachian chorales, an instance of which is shown in Figure 1 below. However, it performs in a way that is unlikely to reflect processes involved in ‘real world’ composition; it can 'only' harmonise Bach chorales, and does so by methods such as generate-and-test that do not seem to be intuitively plausible or to relate to known psychological processes. In particular, it does not offer a model for understanding how Bach might have composed and harmonised chorales; there is no attempt to model the historical provenance of the melody-

---

[1] K. Ebcioglu. 'An Expert System for Harmonizing Chorales in the Style of J. S. Bach', in M. Balaban, K. Ebcioglu and O. Laske (eds.), *Understanding Music with AI: Perspectives on Music Cognition* (Cambridge, Mass: AAAAI Press/MIT Press, 1992), pp. 294-333.

types in the Lutheran tradition so no attempt to set what is produced into, or to guide what is produced by, information about the generic pathways in chorale harmonisation that might have been dependent on the learned style of pre-Bach harmonisations, nor is there an acknowledgement that the preferred harmonisation of a chorale might be governed by its words, nor of the 'embodied' component that would recognise the degree to which the harmonisation 'fell under the fingers'.

Example 1: A 'Bach chorale' written by Kemal Ebcioglu's computer program CHORAL

In effect, what Ebcioglu was proposing was not a model of the workings of the mind in musical composition. It was a powerful demonstration that mechanisms other than human minds can do some of the same things as human minds. Ebcioglu was not really concerned to account for Bach's compositions, but to produce an algorithmic approach to a human problem, that of harmonising chorales in the manner of Bach. He succeeded in this, but the suggestion is not that this is how Bach harmonised but that there might well be a number of possible solutions to the issue of how to harmonise in the manner of Bach, some of which may have computational and information-engineering advantages and possibly generality. Other solutions are required to approach more closely the manner in which a human would undertake the task, and still others might be required to deal with the 'facts', insofar as they are known, of how Bach himself performed this task.

In other words, cognitive science looks in two directions at once, addressing the two domains of computation and cognition, and solutions in one domain might not be solutions in the other. This double viewpoint, which is the result of attempting to model human cognition computationally while trying to offer computational ways of manipulating the objects of the human mind, enables informational engineering to reach solutions that alter the problem by transforming our conception of it through changing the ways we can manipulate it. Informational engineering can do this by finding solutions in the engineering domain to problems in the cognitive domain that transform the nature of these cognitive problems, such as the development of standards of data representation and access that offer new ways of manipulating data and new forms of musical objects and relations, thus affording new ways of interacting with music and with others in music.

Indeed, Ebcioglu's approach to harmonising Bach chorales has something in common with the chess-playing program Deep Blue that in 1996 beat Gary Kasparov. Ebcioglu treats Bach's music as a hypercomplex pattern, and applies complex computational techniques (albeit informed by music theory) to the task of generating it. Deep Blue treated chess as a hypercomplex pattern and applies brute computing power to generate and unravel possible patterns of moves. But Deep Blue does not do this in the same way as a chess grandmaster. It number-crunched phenomenally fast, assessing the consequences of millions of moves in seconds, whereas it appears that chess grandmasters are likely to have a more constrained stock of predictions, based in part on their knowledge of the course of entire games and, perhaps most importantly, on the probable psychology of their opponents. Deep Blue, with no psychology, was a Martian chess-player which could not be 'psyched-out' by the human mind games of its opponents. Similarly, a program such as CHORAL presents a musically intransigent phenomenon that does something that humans can do, but does not necessarily correspond to the way in which humans harmonise chorales.

Today, Ebcioglu's approach, which was developed in the 1980s, would probably not be the first choice to model aspects of music cognition computationally, partly because of its use of explicitly symbolic representations. Many current researchers would reject the symbolic approach, partly because of the difficulty of acquiring the knowledge that it instantiates, and partly because of the explicitness of the mental representations that it appears to assume. Humans acquire knowledge even without explicit instruction, and representations in human minds appear to be implicit rather than explicit.

Hence much recent computational research on music cognition has employed connectionist or distributed systems, in which 'musical knowledge' is not explicitly and atomistically encoded, but is constituted as different emergent states and dispositions of the whole computational model.[2] This 'neural network' approach has had much success in providing functional models of aspects of music cognition. However, this success tends to be focused on more general musical capacities than harmonising a chorale in the style of Bach, such as the capacity to predict and relate sequences of pitches in time so as to represent melodies as linear unfoldings.

The success of many recent connectionist approaches to music has been to model general processes whereby the human mind organises and makes sense of patterns in sound. This focus on characterising general mental perceptual processes, rather than on modelling complex human skills, typifies much of the empirical work that has been conducted on music within cognitive science, and it is to the empirical work that I now turn. In much of this research Bach's music has again been used as a reservoir of complex patterns, but here the prospective derivation of the human capacity to make sense of these patterns from general human capacities to make sense of our sound environment is explored. The most substantial and experimentally grounded theory of how we humans make sense of our sound world has originated with Al Bregman in his theory of Auditory Scene Analysis.[3]

Empirical studies

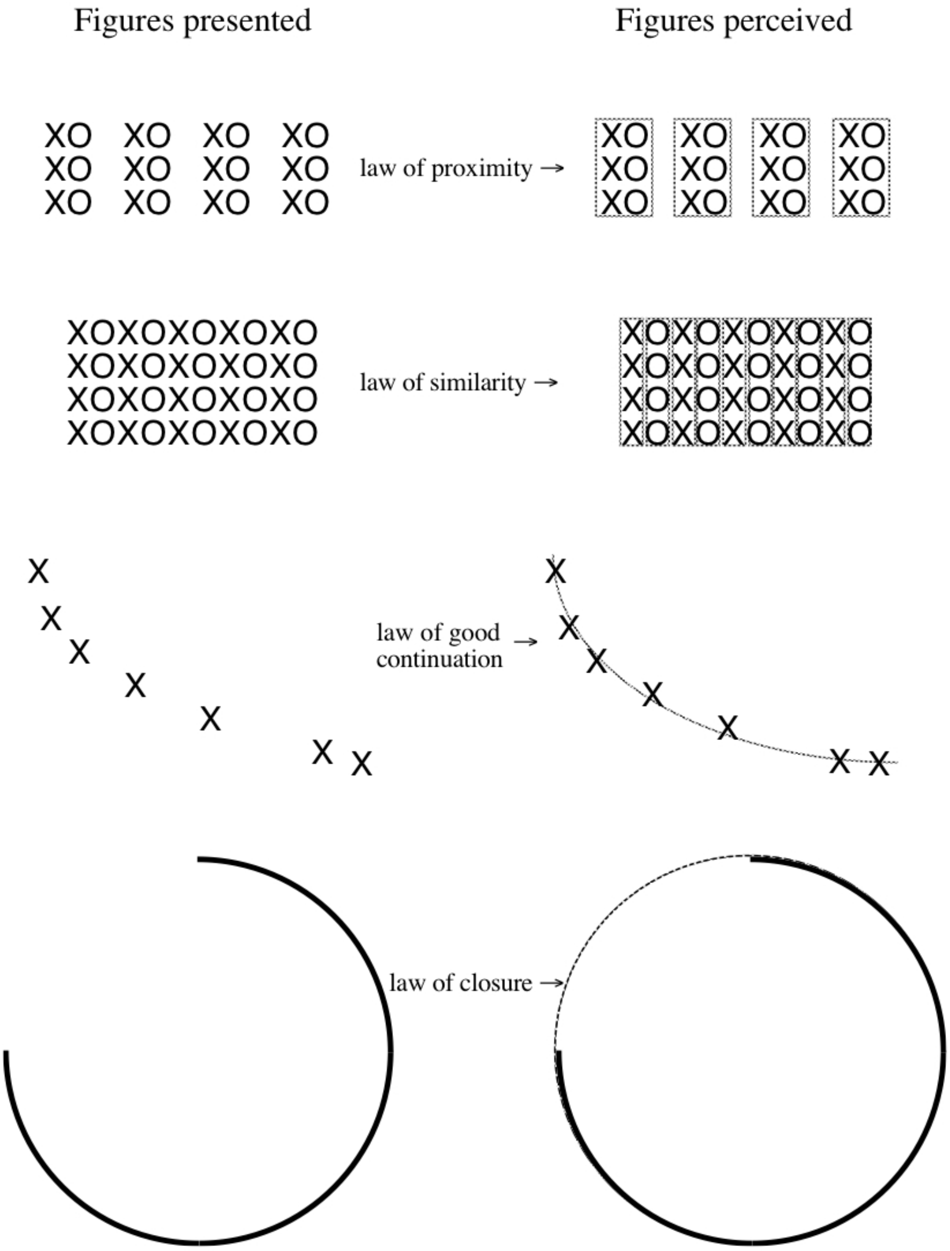
In developing his theory of Auditory Scene Analysis (ASA), Bregman took as his starting point the theories developed by the Gestalt psychologists in the 1920s and 1930s. These theories proposed laws that governed the perception of form, laws such as 'similarity', proximity', 'good continuation' or 'closure', illustrated graphically in Example 2. The Gestaltists interpreted the operation of these laws as deriving from the operation of 'fields' in the brain; however, Bregman reinterpreted these laws as emerging from the ways in which humans notice or infer locations and identities of sources of sound in the auditory environment. These principles arise in part from regularities exhibited by sound-producing objects because of their conformance to the principles of physical acoustics, and in part from the expected as the number of voices occupying the same general pitch range increases.[4] He concludes that it is likely

---

[2] For example, E. R. Miranda (ed.), *Readings in Music and Artificial Intelligence* (London: Routledge, 2000).

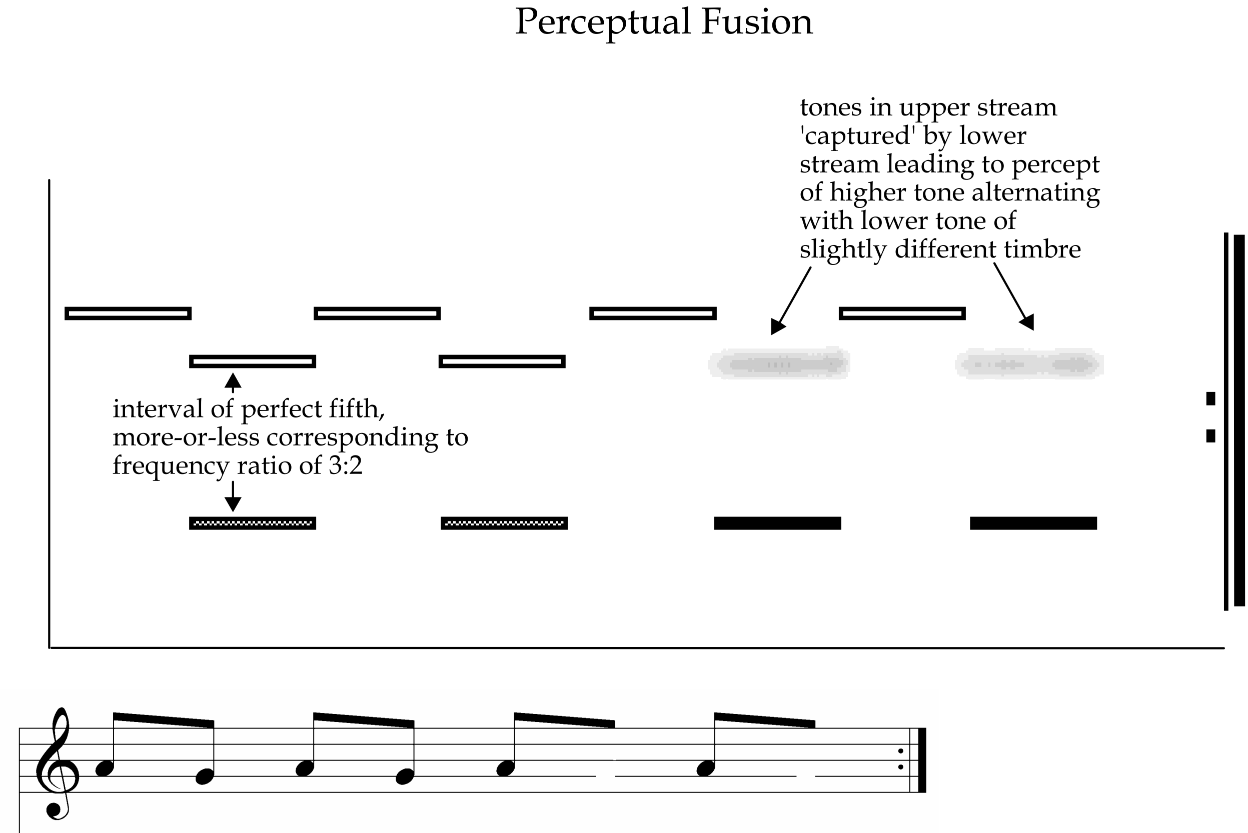
[3] A. S. Bregman, *Auditory Scene Analysis: the Perceptual Organisation of Sound* (Cambridge, Mass.: MIT Press, 1990).

[4] D. Huron. 'The Avoidance of Part-crossing in Polyphonic Music – Perceptual Evidence and Musical Practice', *Music Perception* 9 (1991), 93-104.



Example 2: Visual representations of figural forms arising from the operation of Gestalt processes in perception

that Bach was seeking 'to minimise perceptual confusion as the density of auditory images increased'. In other words, the disposition of voices in these multi-voice works is structured so as to ensure that: 'sounds that have similar properties' [here, that occupy discrete pitch ranges] 'are linked together in groups that are segregated one from another'. Bach here seems to be concerned with *maintaining the independence in perception of the different voices*, and the way in which he does so conforms to basic principles of auditory scene analysis. Similarly, an analysis of a smaller corpus (fifteen polyphonic works) shows that synchronous voice onsets occur statistically significantly less than would be expected by chance. Again, Bach appears to be exploiting an ASA principle, this time suggesting that 'unrelated sounds seldom start or stop at exactly the same time' in order to *ensure the independence of individual voices in perception*. further study in which Huron explored more musically subtle aspects of Bach's polyphonic practices showed that the frequency of occurrence of vertical intervals was negatively correlated with the tendency of intervals to 'promote tonal fusion'. [5] This finding and its significance may require a little explanation.



Perceptual Fusion

tones in upper stream 'captured' by lower stream leading to percept of higher tone alternating with lower tone of slightly different timbre

interval of perfect fifth, more-or-less corresponding to frequency ratio of 3:2

Example 3: Representation of Bregman and Doehring's (1984) findings; the upper part of the figure illustrates the percept that is likely to be experienced, shown in musical notation in the lower part.

In 1984, Bregman and Doehring found that when two simple lines or 'streams' of tones were sounded together in which certain tones in the upper stream stood in low-integer frequency ratios to tones in the lower stream, the low-

---

[5] D. Huron, 'Tonal Consonance Versus Tonal Fusion in Polyphonic Sonorities', *Music Perception* 9 (1991), 135-54.

integer related tones in the upper stream could be perceived being 'captured', or integrated into, the tones of the lower stream, with the resulting percept changing from two simple 'streams' into one stream in which a higher tone alternates with a much lower tone of slightly different timbre.[6] The effect is represented graphically in Figure 3, in which the musical notation represents the stimulus presented, while the upper figure illustrates the percept that is likely to be experienced.

Intervals that are likely to promote 'fusion' are those of the fifth and the octave, which, even in the equal-tempered scale, approximate closely to low-integer ratios (of 3:2 and 2:1 respectively). The tendency to avoid those intervals as verticalities in contrapuntal works, and thus minimise the likelihood that a note in one voice will be 'captured' by a tone in another voice, is in line with an objective of *ensuring that the perceptual independence of the contrapuntal voices is maintained*.

Huron's research has thrown up some results that are not predictable from ASA principles, but that suggest that other, still general, constraints on perception may play a role in shaping the nature of our experience of Bachian counterpoint. In a pair of studies, one experimental and one statistical,[7] it was shown that listeners tended to become confused in identifying voice-entries in a polyphonic texture when more than three voices were present, particularly where those entries were in inner voices, and that Bach appears to have sought to minimise the potential for confusion by avoiding inner-voice entries in textures of five voices. Part of the significance of this experiment is that both highly trained musicians and untrained listeners participated, and although the musicians performed with greater acuity than the untrained participants, the results of both groups showed the same pattern of confusions. It seems that when confronted with complex polyphonic textures, regardless of their level of musical training, listeners' capacity to assess the number of voices in the texture conforms to the model 'one, two, three, many…'; it also seems that Bach was well aware of this; but perhaps the fact that he does not avoid inner-voice entries in four-voice textures suggests that he might have felt that we really should be able to do better.

Huron is by no means the sole researcher in this field, and in his recent summative paper he integrates the results of his previous research with a broad range of other psychoacoustic and psychological research to produce a general theory of voice-leading that is grounded in perceptual principles.[8]

---

[6] A. S. Bregman and P. Doehring, 'Fusion of Simultaneous Tonal Glides: The Role of Parallelness and Simple Frequency Relations', *Perception & Psychophysics* 36 (1984), 251-56.

[7] D. Huron, 'Voice Denumerability in Polyphonic Music of Homogeneous Timbres', *Music Perception* 6 (4), 1989, 361-82 and D. Huronand D. Fantini, 'The Avoidance of Inner-voice Entries: Perceptual Evidence and Musical Practice' *Music Perception* 7 (1), 1989, 43-48, respectively.

[8] D. Huron, 'Tone and Voice: A Derivation of the Rules of Voice-leading from Perceptual Principles', *Music Perception* 19 (2001), 1-94. The paper is available online at: http://dactyl.som.ohio-state.edu/Huron/Publications/huron.voice.leading.html).

Future directions

Taking these two examples, we must ask what the cognitive science of music has contributed to our understanding of the music of J. S. Bach. It tells us that processes quite unlike those which Bach probably used can generate music not dissimilar from the music that Bach wrote. It tells us that our experience of Bach's music is mediated and shaped by general perceptual processes, and that Bach's shaping of his own music is likely to have been mediated by those same processes. But so far the cognitive sciences have tended to address music either as notation, or as the human response to organised sound, while apparently taking little account of how humans organised sound and action in music. If the musicology of the last thirty years has demonstrated anything at all with clarity, it is that music is a property of human collectivity, of human cultures, if you will, rather than of the individual, and the cognitive sciences have to date tended to focus on the individual capacity for, and response, to music.

This could be presented as an apparent disjunction between music as a biological phenomenon and music as a cultural phenomenon.[9] Much of the cognitive science of music appears to regard music as a biological phenomenon. Researchers in cognitive science must recognise the existence of this disjunction, even if they cannot bridge it, and work out strategies for broadening cognitive science's scope beyond the 'biological'. Future researchers will need to read, listen and even talk to musicologists more than has previously been the case. Of course this cuts both ways, and musicologists too must be willing to engage with the concerns of the cognitive sciences. Three possible areas stand out as the focus of future research, and all three are likely to demand broader musicological underpinnings than previous areas. These three areas are music as interaction, music as embodied action, and music as embedded in other domains of human life.[10]

Biological as well as cultural phenomena are probably better conceived of in terms of interactions than in terms of individual actions. Music is no exception; it can be regarded as constituted in performance through the interactions between the performers and between performers and audience.[11] We now have methods and computational tools that enable us to tackle the issue of understanding musical interaction and that might afford the same degree of quantitative rigour that we can apply to individuated experience. The nature

---

[9] I. Cross, 'Music as Biocultural Phenomenon', *Annals of the New York Academy of Sciences (The Neurosciences and Music)*, 2003: Vol 999, G. Avanzini, C. Faienza, D. Minciacchi, and others (eds.), 106-11.

[10] I. Cross, 'Music and Biocultural Evolution' in M. Clayton, T. Herbert and R. Middleton (eds.), *The Cultural Study of Music: a Critical Introduction* (pp. 19-30) (London: Routledge, 2003), pp. 19-30.

[11] For example, P. Juslin, 'Communicating Emotion in Music Performance' in P. Juslin and J. A. Sloboda (eds.), *Music and Emotion: Theory and Research* (Oxford: OUP, 2001) pp. 309-37.

and quality of the interactions that constitute music must be explored in any cognitive modelling of music, and be examined as causal elements in any experimental investigation of music cognition.

Cognitive science must recognise that music is evidenced in action as well as sound. While a considerable amount of research has focused on this perspective in the context of the study of performance,[12] we also need to consider how it impacts our experience as listeners and perhaps as composers. The cognitive neuroscience of music, an area I have not considered here, suggests that even in listening to music, areas of the central nervous system associated with movement and motor control are activated, even for those without formal performance experience.[13] It seems that our listening experiences are indeed modulated by the dynamics of our embodiment.

Cognitive science must also recognise that music is embedded in other domains of human life. For example, I have already mentioned the idea that for Bach, the words to which a chorale was set were a powerful determinant of the setting. Cognitive science can shed little light on this at present, although a recent study by Zbikowski presents some intriguing prospects for approaching the issues of text-music relations from a cognitivist perspective.[14] Cognitive science needs to address the ways in which music in cognition relates to, and is bound by, the cognitive processes that constitute other domains of human experience such as language, emotion, and processes of learning and development.

Promising work within the cognitive sciences of music is already being conducted in these three domains. The future is exciting, but there is still a great deal to learn about the experience of musical structure in the abstract. Indeed, the notion of Bach's music as hypercomplex pattern still has appeal.[15] The traditional methods and concerns of the cognitive sciences have by no means been exhausted, and it might just be that Bach's music will in the end turn out to be explicable in terms of some set of formal principles.

---

[12] For example, A. Gabrielsson. 'The Performance of Music' in D. Deutsch (ed.), *The Psychology of Music* (2nd edn London: Academic Press, 1999) pp. 501-602.

[13] P. Janata and S. T. Grafton, 'Swinging in the Brain: Shared Neural Substrates for Behaviors Related to Sequencing and Music', *Nature Neuroscience,* 6 (7), 682-87.

[14] L. M. Zbikowski, *Conceptualizing Music: Cognitive Structure, Theory and Analysis* (Oxford: OUP, 2002).

[15] See, for example, M. Rohrmeier, *Towards modelling harmonic movement in music: Analysing properties and dynamic aspects of pc set sequences in Bach's chorales*. (M.Phil thesis, University of Cambridge, 2006): available at http://www.dar.cam.ac.uk/dcrr/dcrr004.pdf