Extra-Music(ologic)al Models for Algorithmic Composition.

No Author Given

No Institute Given

Abstract. It is proposed in this paper that in designing algorithmic composition systems it may be useful to consider the important characteristics of music from the listeners' perspective, rather than the analysts. In doing so the fundamental dynamic nature of music, which is lost in transcription, becomes apparent. Initial explorations of dynamical systems are presented as a means of producing a sense of movement, and creating higher level structures.

1 Introduction

There is now a large interdisciplinary community of researchers and practitioners working in the field of automated composition, with very different motivations, inspirations and techniques. A few years ago Pearce et al [16] suggested that the field suffered what they termed a 'methodological malaise', due to a pervasive failure to specify the motivations and goals of research. They outlined four different activities, with different motivations, and proposed suitable methods of evaluation. These were:

– **Algorithmic composition** for expanding the compositional repertoire.
– **Developing compositional tools** for use by composers.
– **Musicological modelling musical styles** to evaluate stylistic theories.
– **Cognitive modelling of the composition process** to evaluate cognitive theories.

The primary focus here is on Algorithmic composition, with the aim of producing music to 'expand the compositional repertoire for human listeners', [16]. This can be most sensibly assessed, as suggested, in the same way that all other music is appraised: through listeners' response to public performance and/or publicly distributed recordings. This is a seemingly obvious, but important point as it implicitly states that our primary concern is development of a system capable of producing music that serves the same function as music created in more traditional ways. We are not concerned with modelling either the compositional process or evaluating stylistic theories of existing work. We are concerned with producing something that people can have some "some degree of meaningful or gratifying perceptual engagement with" [7]. With this in mind, it is proposed in this paper that the design of algorithmic composition systems should not rely too heavily on principles drawn from music theory, and that a potentially profitable approach may be to consider more closely the functional characteristics of music from an experiential, rather than analytical perspective.

Approaches to Algorithmic composition are commonly categorised according to the principles techniques used in their construction eg [23], [15], however they can also be differentiated according to the extent to which they draw from existing music.

For those familiar with the development of computer models, Western Musicology offers an attractive body of theory which can be embodied as a rule-base or set of constraints. As well as explicit encoding in Knowledge Based systems eg [9], theoretic principles can be embodied more implicitly in evolutionary or learning algorithms. For example the rules of harmony offer an apparently neat solution to the headache of designing a fitness function for a Genetic Algorithm; spatial representations of tonal distance eg [18] p.20, [20]) offer a possible error measure for an Neural Network. And indeed this is common practice: the "Fitness function judges the fitness of each chromosome according to criterion taken directly from music theory" [17].

At the other extreme, some practitioners employ mathematical models which are selected on the basis that they represent an abstract model of some musical phenomena. These are often described as an "extra-musical" because "their 'knowledge' about music is not derived from humans or human works" [15] p 2. Although there are some exciting examples of extra-musical applications such as Blackwell's Swarm system [3] many see this approach as little more than an interesting novelty and even those who have adopted this practice suggest that the success in terms of human appeal is limited, as Miranda puts it: "the results normally sound very strange to us" [14] p.1. Perhaps because of the idiosyncratic nature, these systems are considered to be inherently less 'musical'.

2 Musicological vs Musical Perspectives.

2.1 Music for the Analyst.

Music theories in their many guises, represent attempts to *understand* music in a Musicological sense; the analyses from which rules are derived aim to achieve possible coherent sets of principles and ideas with which to rationalise, analyse and investigate the structurally functional aspects of music. This is neither exhaustive, nor aimed primarily at describing music in terms of the listener's perception. "Each musical culture rationalises only a few selected aspects of its musical production ... [so] any cultural representation of music (ie music theory) must constitute a thoroughly incomplete specification of the intended musical experience... A formal analysis is a kind of mechanism whose input is the score, and whose output is a determination of coherence...In other words, it purports to establish or explain what is significant in music while circumventing the human experience through which such significance is constituted; to borrow a phrase from Coulter, it aims at 'deleting the subject'" [7] p.241

Cook's basic argument is that there is an important and inevitable discrepancy between the experience of music aurally, and the ways in which it is imagined or thought about. He draws a useful distinction between 'musical listening' which is concerend with the aesthetic gratification in being absorbed in a non-dualistic sense, and 'musicological listening' for the purpose of establishing facts of the formulation of theories about the music, perceptual object. This is not to say that the thoughts and ideas of theorists and analysts are inadequate or misplaced, but simply that they are not aimed primarily at giving a perceptual account of music.

An extreme example of this discrepancy between analytic and experiential reality is provided by experiments in which two versions of short piano pieces were played to music students: their original form, which began and ended in the same key, and an altered form which had been modified so as to modulate to, and end, in a different and unrelated key [5]. In standard music theory, tonal closure - or more generally the influence, or organising function of the overall tonic - is the very core of the traditional forms of eighteenth and nineteenth century music. However in these trials, there were no statistically significant differences in preference for the original over the altered forms, even though in many cases the modifications meant the pieces ended up in keys as distant as the minor second. In another set of tests [6], music students who were played the first movement of Beethoven's G major sonata (Op 49 No 2) frequently predicted that the music would carry on for another minute or more when the performance was broken off just before the final two chords. Theoretically, the recapitulation and coda are key functional structures, signifying the close of a piece, but here seem to be utterly ineffectual when presented aurally.

Perhaps more elegantly designed studies would be needed to make any strong claims, but it is common for musicologists to differentiate between the aural and analytic aspects of a piece of music. Kathryn Bailey writes of Webern's symphony that it consists of "two quite different pieces - a visual, intellectual piece and an aural, immediate piece, one for the analyst and another for the listener" [2] p.195. Thomas Clifton expresses this more incisively: "For the listener, musical grammar and syntax amount to no more than wax in his ears". [4] p.71.

Discrepancies between the experientially and theoretically fundamental properties of music arguably stem from the discrepancy between the nature of time as it exists in aural and written music. The contradiction between the ever-present 'inner' time in which music is experienced, and the retrospective 'outer' time which is imposed in the act of reflection and measured by musical notation is as a fundamental dilemma for many theorists. Schutz suggests that attempts to describe the musical experience in 'outer' time poses a variant of the Eleatic paradox - ie that the flight of Zeno's arrow cannot be described because it is impossible to represent the ongoing quality of its' motion. As Schutz puts it "you may designate the spot occupied by the arrow at any chosen constant during the flight. But then you have dropped entirely the idea of an ongoing motion." [19] p.30. At the heart of this lies the discrepancy between the static, symbolic nature of music in notated form - which is the principle object of music theoretic concern - and the dynamic immediate nature of music in sonic form - which is the object of the listener's concern.

Many composers have attempted to focus on the aspects of the music that *are* most relevant to listeners perceptual experience eg [25], [22]. Some of the ideas of Ernst Toch are presented in following section to illustrate an alternative means of understanding the functional aspects of music from the listeners, rather than analyst's perspective.

2.2 Music for the Listener

> "I never expected so much fascination to come from investigations of the nature of musical theory and composition. Aspects unfolding to me show why the rules of established musical theories could not be applied to 'modern' music, why there seemed to be a break all along the line, either discrediting our contemporary work or everything that has been derived from the past. To my amazement I find that those theories are only false with reference to contemporary music because they are false with reference to the old music, from which they have been deduced; and that in correcting them to precision you get the whole immense structure of music in your focus."' [22] p xii

Ernst Toch was a masterful and original classical composer, renowned also for his paramount studio film scores. Later in life he became preoccupied with the reconciliation of theories of classical music with contemporary modernist trends. *The shaping forces in Music* is his account of how all musical writing must respond to the psychological wants of the listener, and how similar goals may be achieved in different styles. If harmonic structure is the cornerstone of traditional music theory, Toch sees the movement of melodic 'impulses' as the central force of music from the listener's perspective. He describes harmony as 'arrested motion' by which he means to stress the fundamental Heraclitean flux in music.

Harmony as Arrested Motion
In an example that is typical of the tasks used in algorithmic approaches to harmonisation Toch presents a phrase from a folk tune, that invites a simple I, IV V harmonisation (Fig 1 A). This is something that a Genetic Algorithm could perhaps achieve. We could even potentially implement theoretic axioms for finding appropriate chord inversions: by minimising the number of steps that each note must take into membership of adjacent harmonies we could feasibly find the first inversions needed to create a smoother chordal structure in bars 2 and 3 as shown in Fig. 1 B. The apparent simplicity and efficacy of this kind of 'rule' is precisely what is attractive to the algorithmic composer, but as Toch warns: "While this axiom seems a simple expedient for the beginner, it implants in him a dangerous misconception, namely the view point of rigidly preconceived harmony as a fixed unit, within the frame of which each voice seeks to take up its' appropriate place." [22] p5.

This point is illustrated by considering a common or garden Chorale harmonisation shown in Fig. 2 A., which concedes to all the traditional rules of harmony. Toch then offers 12 other possibilities, examples of which are given in Fig.2 B and C, which he arrives at by a more general principle which he calls 'linear

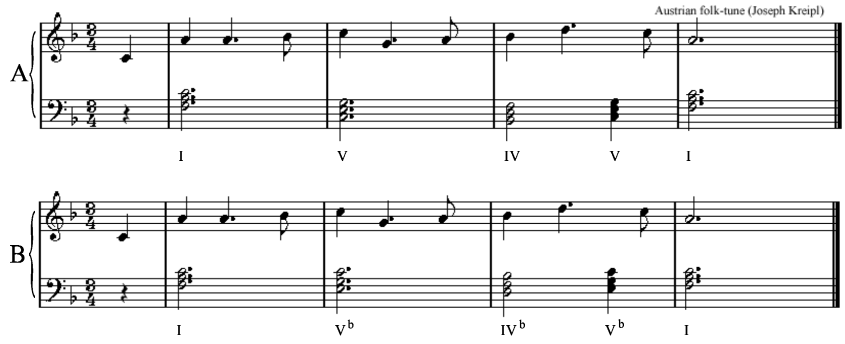

Fig. 1. A Natural harmonisation of a phrase from a simple folk tune using I (tonic), V (dominant) and IV (sub-dominant) (top) and appropriate chordal inversions (IV^b etc)(bottom)

Fig. 2. Standard harmonisation of a phrase from a Chorale (A) and examples from Toch's alternative harmonisations (B) and (C)

voice leading' - a term he uses to describe the dynamic impulse of each voice. In contrast to the 'appropriate' harmonisations of Fig. 2 A, some harmonies in Fig. 2 B and C go against every rule in the book: consecutive fifths, cross relations, arbitrary dissonances etc. "And yet we hope that the reader, even though these harmonisations may appear unusual and strange, will feel their logic and organic life.[1] That they are arrived at by the movement of melodically independent voices is obvious. The truth is that the melodic impulse is primary, and always preponderates over the harmonic; that the melodic, or linear impulse is the force out of which germinates not only harmony but also counterpoint and form. For the linear impulse is activated by *motion* and motion means life, creation, propagation and formation". ibid p.10.

Toch's point here, is that harmonies are not dictatorial pillars which define the pitch of the constituent notes, but snapshots of coincidences between separate melodic lines as they develop in time. This stands in stark contrast to the way in which the 'harmonisation problem' is sometimes conceived and approached in algorithmic composition: "We apply the following criteria: we avoid parallel fifths, we avoid hidden unison, we forbid progression from diminished 5th to perfect 5th; we forbid crossing voices ... From an aesthetic perspective, the results are far from ideal: the harmonisation produced by the GA has neither clear plan or intention" [17] p.5.

This is an extreme, although not atypical application of music theory to the design of algorithmic systems. Considered in the light of Cook's comments on the nature of musicological listening, and Toch's comments on the nature of musical listening some potential problems with this approach come to light. Music theory, working with a static representation of music, forms abstractions and generalisations. As Cook suggests, some of the key functional structures maybe imperceptable aurally. It seems far from inevitable that in using theoretic principles as guiding principles to design systems capable of creating and playing music that we will recreate the fundamentally dynamic surface structure. And without this, it may be hard to produce a sense of plan or intention that underlies the structures from which theory generalises. This could be one reason why we frequently see comments such as: "while conforming to classical triadic harmony, the music seems lifeless" [8] p.21 or "The music often wanders with unbalanced and uncharacteristic phrase length. No musical logic is present beyond the chord-to-chord syntax" ibid p.22

[1] Toch invites the reader to play each line, separately at first and then with the soprano, before playing the full harmonies, listening to each separately to appreciate their movements. For those with no piano to hand examples can be found on-line at http://www.cogs.susx.ac.uk/alicee/users/toch

3 Adaptive Systems Music: Exploring Mechanisms for Creating Novel Progressions.

In an ongoing project, possible mechanisms for producing a sense of coherent movement in polyphonic systems are being explored using various networks of dynamical systems. Any model based on differential equations offers a means of representing the temporal flow of musical parameters, as the system is defined in terms of change in state over time. In a sense, the Eleatic paradox is resolved. Both the systems described below are formed of individual nodes which are assembled in small networks. In each case the state of each node is a function (directly or indirectly) of the activity in the rest of the network. An attractive property of this class of dynamical system then, is that as well as enabling the representation of dynamical structures, a form of logical structure can be created. The numerical output of each node is mapped onto frequency deviations. In the examples given here, each network comprises four nodes, so creates four part polyphonic output. This means that as well as describing the 'melodic' evolution of individual parts there is also a certain logical relationship between each part. An interesting possibility being examined, is whether these properties can be transformed into a musically coherent relationships.

3.1 Neural Oscillator networks

Neural Oscillators have been used in robotics tasks requiring rhythmic movement such as sawing [24], and drumming [12], and in models of rhythmic entrainment [21]. Here, a small network of simple neural oscillator model as described by [13], was built.

The oscillator system consists of two simulated neurons arranged in mutual inhibition, as shown in figure 3. The time evolution of the oscillator is given by (1 to 5), where $[x]^+ = \max(x, 0)$. The output of the oscillator is y_{out}, β and γ are constants (here set to 2.5), c is a constant that determines the amplitude of the oscillation and $\tau 1$ and $\tau 2$ are the time constants that determine the natural frequency (in the absence of input), and shape of the output signal. Inputs (g_j) to the oscillator are weighted by gains h_j.

$$\tau_1 \dot{x}_1 = c - x_1 - \beta v_1 - \gamma [x_2]^+ - \sum {}_j h_j [g_j]^+ \ . \quad (1)$$

$$\tau_2 \dot{v}_1 = [x_1]^+ - v_1 \ . \quad (2)$$

$$\tau_1 \dot{x}_2 = c - x_2 - \beta v_2 - \gamma [x_1]^+ - \sum {}_j h_j [g_j]^+ \ . \quad (3)$$

$$\tau_2 \dot{v}_2 = [x_2]^+ - v_2 \ . \quad (4)$$

$$y_{out} = [x_1]^+ - [x_2]^+ \ . \quad (5)$$

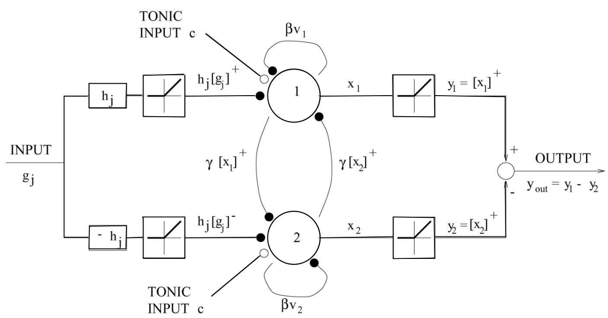

Fig. 3. Schematic of a neural oscillator node. The oscillator equations simulate two neurons in mutual inhibition as shown here. Black circles correspond to inhibitory connections, open to excitatory. The mutual inhibition is through the $\gamma[x_i]^+$ connections ($[x]^+ = \max(x, 0)$), and the βv_i connections correspond to self-inhibition. The input g_j is weighted by a gain h_j, and then split into positive and negative parts. The positive part inhibits neuron 1, and the negative part neuron 2. The output of each neuron y_i is taken to be the positive part of the firing rate x_i, and the output of the oscillator as a whole is the difference of the two outputs.

If an oscillatory input is applied, the node will entrain the input frequency - producing an output of equal frequency, but not necessarily the same phase, as the input. This can be shown to be true over a wide range of input amplitudes and frequencies.
The nodes are arranged in a network such that the input signal to one pair is the output from one or more other pairs. This creates a collection of continuous periodic output signals, with differing phase and form which are synchronised. These outputs are then mapped onto frequency, either discrete or continuous. Although not restricted to any particular key, the periodic form - similar to the wave-like structure of many melodies [22]- produces a sense of movement. Because all outputs are synchronised, but exhibit different structures and phases, a sense of ensemble is achieved with voices moving in unison or opposition according to phase. Example output from a simple network can be found at: http//www.cogs.susx.ac.uk/users/alicee/nosc. In this current form, the outputs - being periodic - are very repetitive. This is not presented as music in itself, but a possible mechanism for conveying movement, and relationships between parts.

3.2 Homeostatic Harmonies

In previous work, the potential for generating harmonies using principles of homeostasis has been explored. A model based on Ashby's description of his electromechanical homeostat [1] was used. The system can be conceived as a number of interconnected nodes, with varying degrees of connectivity. A fully

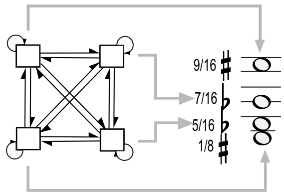

Fig. 4. Schematic of a fully connected four unit homeostatic network (left) and mapping of output values to microtonal pitch values (right)

connected network of four nodes is shown in Fig. 4. The output of each node is updated according to the weighted sum of the output of all other nodes (6). If the output of any node exceeds a prespecified value, all the weights in the network are re-randomised, so by a process of trial and error, the system converges on a point or periodic attractor. Once stabilised. If perturbation is within this critical limit, the system returns to the previous attractor after a short period of random oscillation, if perturbed beyond this limit, weights are re-randomised until a new stable attractor is found. The output values are then mapped onto continuous frequency variations, the various states of the network producing either repeated polyrhythmic patterns, or periods of more stochastic behaviour before returning to the same pattern, or settling on a new one depending upon the extent of the perturbation. Full details can be found in [10].

$$\mathrm{O}_{i(t+1)} = \sum_{j=0}^{j} I_{ij(t)} \times W_{ij(t)} \text{ where } I_{ij}(t) = \sum_{j=0}^{j-1} O_{j(t-1)} + O_{i(t-1)} \quad . \qquad (6)$$

Where $O_{i(t+1)}$ is the Output of the i_{th} unit at time $_{t+1}$, $I_{ij(t)}$is the input to the i_{th} unit from the j_{th} and $W_{ij(t)}$ is the weight from unit j to unit i.

This basic system has been used both within a ensemble of live musicians, to generate material for a composition commissioned for the LUX open film festival, and in conjunction with a cellular automata as ‘background’ music in an exhibition space at a science art forum[2]. Examples of outputs can be found at http://www.cogs.susx.ac.uk/users/alicee/AdSym. Although not pertaining to any particular musical style or genre, there is a sense of movement and direction in the output. In a listening test, all participants agreed that it was ‘musical’, elaborating on their choice with comments such as: ‘sense of melody’, ‘there were definite harmonies if unusual at times’, ‘sense of harmonic structure and melodic progression’. This reference to structure was made by several listeners: ‘structure and development on different time scales/resolutions’. The further comments of many listeners suggest that the music had emotive qualities: ‘tension building and resolution of tension.’ [10].

[2] http://www.blip.me.uk

The system that these listeners heard also included a Cellular Automata model, the states of which determined *when* the current pitches defined by the homeostat were voiced. It is certainly far from clear what aspect of the system was responsible for the reported experiences of progression and movement etc. However, initial results suggest that the dynamical structures of these algorithms produce a sense of movement in musical terms. Although it can not be claimed that the logical structures of the systems are responsible for reported perceptions of 'harmonic progression' etc., initial experiments using sound to analyse complex systems, suggest that the qualitative state (eg chaotic, complex, ordered) of some systems can be appreciated when presented aurally [11](in press). This makes the development of 'bespoke' extra-musical algorithms for an interesting possibility for generating both surface movement, and larger scale structures.

4 Summary and Discussion.

We have briefly looked at the nature of musicological thinking which underpins the music theoretic constructs commonly used as guiding principles in the development of algorithmic systems and contrasted this with consideration of the perceptually important aspects of music from a composer's perspective. Whilst musicological research may be vital for designing automated composition systems to asses stylistic theories of music, or even cognitive theories of traditional composition (the act of transcription arguably being a vital part of the creative process), we argue that the development of algorithmic systems aimed at engaging the listener may benefit from greater consideration of the perceptually fundamental aspects of music. The importance of coherent movement is one such characteristic.
Examples of attempts to capture a sense ensemble or harmonic progression as an 'emergent' property of melodic movement using simple dynamical models were given. Although the relations between the nodes does not follow any musical model, it is an interesting possibility that the internal logical coherence creates musical structures that listener's describe as harmonic or melodic progressions. If we do wish to create music within the 20th century tonal tradition it is feasible that this approach could be combined with tonal concepts, for example by controlling the mapping of the outputs onto frequency values that accorded with a desired key, or evolving the parameters of the models to achieve the desired dynamical properties. However, if we wish to truly wish to *expand* the repertoire, the development of dynamical systems of this class, in conjunction with evolutionary approaches for parameter tuning offers an intriguing method for exploring the possibility of novel engaging or meaningful musical structures.

References

1. W.R. Ashby. *Design for a Brain; The Origin of Adaptive Behavior.* Chapman and Hall Ltd and Science Paperbacks, 1965.

2. K. Bailey. Webern's opus 21: creativity in tradition. *Journal of Musicology*, 2:184–195, 1983.
3. Young M. Blackwell, T. Swarm granulator. In *EvoWorkshops Proceedings*, volume 3005 of *Lecture Notes in Computer Science*. Springer, 2004.
4. T. Clifton. An application of goethe's concept of steigerung to the morphology of diminution. *Journal of music Theory*, 14:165–189, 1970.
5. N. Cook. Musical form and the listener. *Journal of Aesthetics and Art Criticism*, 46:23–29, 1987.
6. N. Cook. Perception of large scale tonal structure. *Music Peception*, 5:197–205, 1987.
7. N. Cook. *Music, Imagination and Culture*. Clarendon, Oxford, 1990.
8. D. Cope. Computer modeling of musical intelligence in emi. *Computer Music*, 16(2):69–83, 1992.
9. K. Ebcioglu. An expert system for harmonizing four-part chorales. *Computer Music Journal*, 12(1):43–51, 1988.
10. A.C. Eldridge. Adaptive systems music: Algorithmic process as musical form. In *Proceedings of the 2002 Generative Art Conference*, Milan, December 2002.
11. A.C. Eldridge. Listen to your system: The potential for auditory display in artificial life research. *Artificial Life Special Issue: Visualisation Techniques*, 2005.
12. S. Kotosaka, S.; Schaal. Synchronized robot drumming by neural oscillator. *Journal of the Robotics Society of Japan*, 19(1):116–123, 2001.
13. K. Matsuoko. Sustained oscillations generated by mutually inhibiting neurons with adaptation. *Biological Cybernetics*, 52:367:376, 1985.
14. E.R. Miranda. On making music with artificial life models. In *Consciousness Reframed 2003*, 2003.
15. G. Papadopolous and G.A. Wiggins. Ai methods for algorithmic composition: a survey, a critical review and future prospects. In *Proceedings of the AISB '99 Symposium on Musical Creativity*, pages 110–117, Brighton, UK, 1999. SSAISB.
16. M. Pearce, D. Meredith, and G. Wiggins. Motivations and methodologies for automation of the compositional process. *Musicae Scientiae*, 2002.
17. A. Phon-Amnuaisuk, S. Tuson and G. Wiggins. Evolving musical harmonisation. In *Proceedings of ICANNGA*, 1999.
18. A Schoenberg. *Structural Functions of Harmony*. London, 1969.
19. J.R. Schutz. Fragments of the phenomenology of music. In J. Smith, editor, *In Search of Musical Method*, pages 5–71. London, 1982.
20. R.N. Shepard. Structural representations of musical pitch. In D. Deutsch, editor, *The Psychology of Music*, pages 343–390. Academic Press, New York, 1982.
21. M.H. Thaut. Neural basis of rhythmic timing networks in the human brain. *Annals of the New York Academy of Sciences*, 999:364–373, 2003.
22. E. Toch. *The Shaping Forces of Music*. Dover Publications, New York, 1948.
23. P.M Todd and G.M. Werner. Frankensteinian methods for evolutionary music composition. In N. Griffith and P.M Todd, editors, *Musical Networks: Parallel Distributed Perception and Performance*, pages 173–190. MIT Press, Cambridge, MA, 1998.
24. M. Williamson. Oscillators and crank-turning: Exploiting natural dynamics with a humanoid robot arm. Technical report, Hewlett Packard, 202-132, 2002.
25. T. Wishart. *On Sonic Art*. Imagineering Press, York, 1985.