10. What makes music 'music'? Theoretical explorations using zygonic theory

Adam OCKELFORD

1. Introduction

It seems that the great majority of people, across the world, engage with music in one way or another, consciously or unwittingly, every day of their lives. We take it for granted that, quite intuitively, and without the need for formally acquired knowledge or skills, music makes sense to us and has the power to communicate on an emotional level. Although the functions of music vary from one culture to another (Merriam, 1964; Nettl, 2005), and while in some societies music is not conceptually discrete from other forms of social activity, music does have the capacity to exist as pure sound, with no external referents (as in certain Western classical genres, for example). In these circumstances music has no tangible meaning. How, then, does it 'work'? How is it that, just by listening, music offers a coherent aesthetic experience? We will approach this question first by examining the same issue in relation to language-based art forms, which, as a reflection of an external 'reality' or potential, have a more evident source of meaning construction.

2. Meaning construction in verbal language

According to Eliot (1920/1997; 1933), literature has three principal sources of aesthetic response:

- an **objective correlative** (a set of objects, a situation, a chain of events, which shall be the 'formula of a particular emotion');
- the **manner of representation** (including, for example, the use of metaphor);
- the **sound qualities** and **structure** of the language itself.

This thinking may be represented as shown in figure 1. In semiotic terms, the model captures the stages corresponding to the transition from:

- **semantics** (the relationships between signs and the things to which they refer); through
- **syntactics** (the relationships between signs); to
- **pragmatics** (the relationships between the signs and the effects they have on readers or listeners).

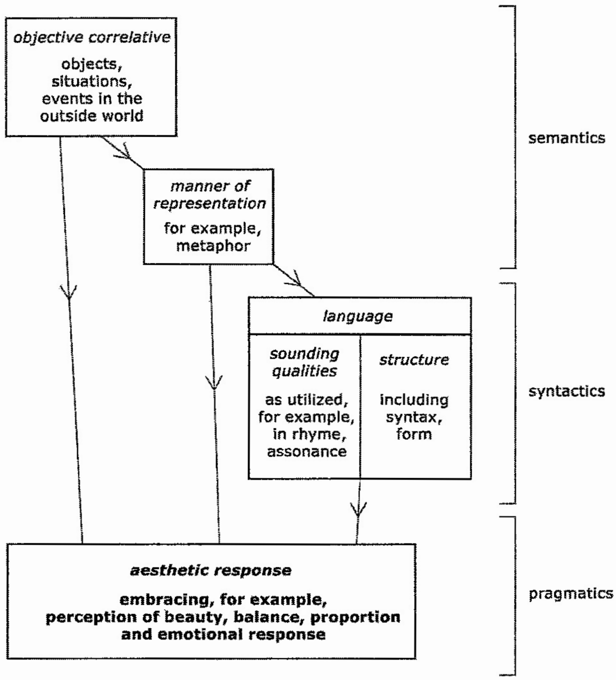

Figure 1 – Representation of Eliot's model of aesthetic response to literary works, and its correspondence to semiotic thinking.

3. Meaning construction in music

However, absolute music (and the abstract component of music that has referential import) has no objective correlative (fig. 2). In these circumstances, how is meaning constructed and conveyed? How is an aesthetic response achieved?

I hypothesise that *all* sounds and the relationships we perceive between them can potentially cause or enable an emotional response (*cf.* Johnson-Laird & Oatley, 1992; Sparshott, 1994, p. 28). There appear to be two main sources of such responses (a) 'expressive non-verbal vocalisations' and (b) 'music-specific' qualities of sound.

'Expressive non-verbal vocalisations' comprise the cues used to express emotions vocally in non-verbal communication and speech (Juslin, Friberg, & Bresin, 2001-2002). They are present cross-culturally (Scherer, Banse, & Wallbott, 2001), suggesting a common **phylogenetic** derivation from 'non-verbal affect vocalisations' (Scherer, 1991) and apparently embedded **ontogenetically** in early maternal/infant interaction (Malloch, 1999-2000; Trehub & Nakata, 2001-2002). It seems that these cues can be transferred in a general way to music, and music-psychological work from the last 70 years or so has shown that features such as register, tempo and dynamic level do relate with some consistency to particular emotional states (Gabrielsson & Lindström, 2001). For example, passages in a high register can feel exciting (Watson, 1942) or exhibit potency (Scherer & Oshinsky, 1977), whereas series of low notes are more likely to promote solemnity or to be perceived as serious (Watson, 1942). A fast tempo will tend to induce feelings of excitement (Thompson & Robitaille, 1992), in contrast to slow tempi that may connote tranquillity (Gundlach, 1935) or even peace (Balkwill & Thompson, 1999). Loud dynamic levels are held to be exciting (Watson, 1942), triumphant (Gundlach, 1935) or to represent gaiety (Nielzén & Cesarec, 1982), while quiet sounds have been found to express fear, tenderness or grief (Juslin, 1997). Conversely, as Meyer (2001) observes, 'one cannot imagine sadness being portrayed by a fast forte tune played in a high register, or a playful child being depicted by a solemnity of trombones'.

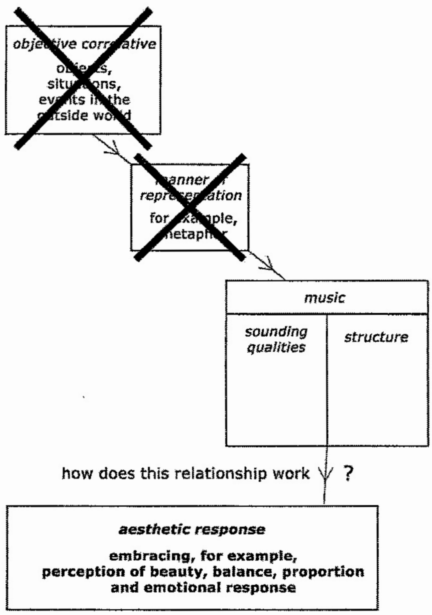

Figure 2 – Absolute music has no objective correlative—so how is meaning conveyed?

'Music-specific' qualities of sound, like those identified above in relation to early vocalisation, have the capacity to induce consistent emotional responses, *within* and sometimes *between* cultures. For example, in the West and elsewhere, music typically utilises a framework of relative pitches with close connections to the harmonic series. These are used idiosyncratically, with context-dependent frequency of occurrence and transition patterns, together yielding the sensation of 'tonality' (Krumhansl, 1997; Peretz, Gagnon & Bouchard, 1998). These frameworks of relative pitch can accommodate different 'modalities', each potentially bearing distinct emotional connotations. In Indian music, for example, the concept of the 'raga' is based on the idea that particular patterns of notes are able to evoke heightened states of emotion (Jairazbhoy, 1971/1995), while in the Western tradition of the last four centuries or so, the 'major mode' is typically associated with happiness and the 'minor mode' with sadness (Hevner, 1936; Crowder, 1985), differences which have been shown to have neurological correlates (Suzuki et al., 2008; Nemoto, Fujimaki & Wang, 2010).

On their own, however, separate emotional responses to a series of individual sounds or clusters would not add up to a coherent musical message—a unified aesthetic response that evolves over time. So what is it that binds these discrete, abstract experiences together to form a cogent musical narrative?

Consider (verbal) language once more. Eliot's 'objective correlative' is likely to be a series of events, actions, feelings or thoughts that are in some way *logically related*, each contingent on another or others through concepts such as causation. Such relationships will be conveyed and given additional layers of meaning through language-specific relationships such as metaphor (in the domain of 'manner of representation'), rhyme and meter (in the domain of 'sounding qualities') and syntax (in the domain of 'structure') (fig. 3).

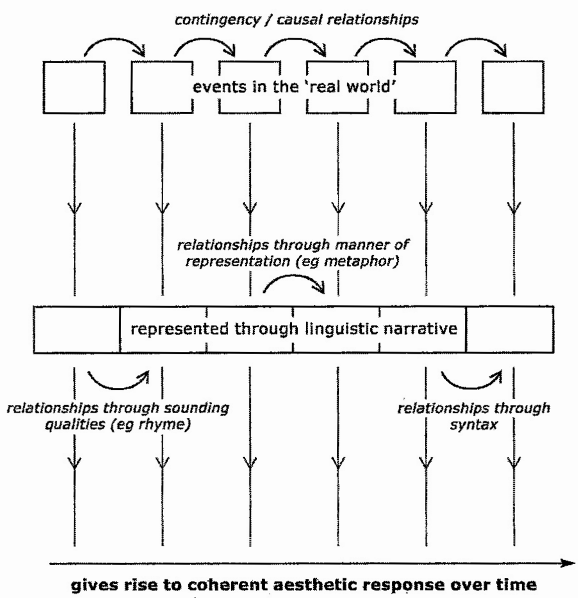

Figure 3 – The forms of logical relationship underpinning meaning in language.

How does a comparable sense of coherence and unity come about in music, when it cannot borrow a sense of contingency from the external world? In the absence of an objective correlative, musical events can refer only to *themselves* (Selincourt, 1920/1956). Self-evidently, one sound does not *cause* another one to happen (it is performers who do that), but one can *imply* another (Meyer, 1989, p. 84 ff.) through a sense of *derivation*. That is, one musical event can be felt to stem from another, and it is my contention that this occurs through *imitation*. If one fragment or feature of music echoes another, then it owes the nature of its existence to its model. And just as certain perceptual qualities of sound are felt to derive from one another, so too, it is hypothesised, are the emotional responses to each. Hence over time a metaphorical (musical) narrative can be built up through abstract patterns of sound.

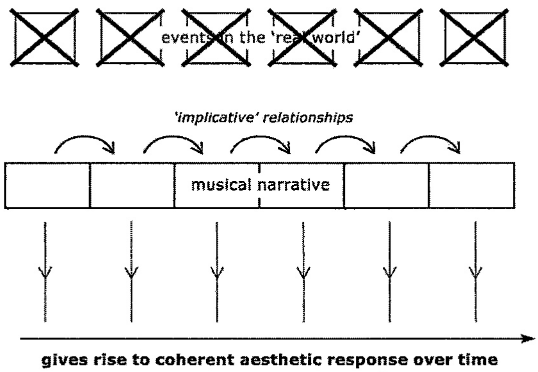

Figure 4 – Relationships underpinning logic in music.

4. Zygonic theory and the importance of repetition in music

This hypothesis lies at the heart of 'zygonic theory' (Ockelford, 1991, 1999, 2005, 2009a, 2011), which predicts that if music makes sense through a feeling of derivation (which stems from imitation), then *repetition* in music should be pervasive. And, indeed, it is: a phenomenon that is recognised in a wide range of literatures. For example, the ethnomusicologist Nettl (2005) finds features common to all musical dialects, including the fact that material is habitually repeated or varied. Indeed, the ubiquity of repetition within pieces is widely recognised. The essayist Selincourt (1920/1956, p. 155) notes that the 'foundation of musical expression is repetition. [It] begins in the bar, and continues in the melody and in every phrase or item into which we can resolve it'. Similarly, the critic and musicologist Zuckerkandl (1956) writes:

> music can never have enough of saying over again what has already been said, not once or twice, but dozens of times; hardly does a section, which consists largely of repetitions, come to an end, before the whole story is happily told all over again. (p. 213)

Composers themselves have expressed the same view. Stravinsky (1942), for instance, observes that 'we instinctively prefer coherence and its quiet strength to the restless powers of dispersion—that is, we prefer the realm of order to the realm of dissimilarity' (p. 69-70); according to Chávez (1961), 'repetition has been the decisive factor in giving shape to music [...] the various devices used to integrate form are, again and again, nothing but methods of repetition' (p. 38 and 41); while the composer and theorist Schoenberg (1967) is characteristically unequivocal: 'Intelligibility in music seems to be impossible without repetition' (p. 20).

Repetition features widely in other theoretical and analytical work too, its presence and functions acknowledged, if not explicitly, then by implication. Consider, for example, the traditional Western notion of form, as espoused by writers ranging from Macpherson (1915) to Berry (1986). Here, the concept of stereotyped structures such as AA'A"A'"... (characteristic of variation sets), ABA ('ternary' form) and ABACA... (the 'rondo') implicates repetition both within pieces and between them. Schenker also acknowledges the part played by repetition, both at the level of motives and in the construction of large-scale forms, in his early *Harmonielehre* of 1906. This recognition carries over into the sophisticated models of musical structure that followed: in *Free Composition* (1935) the question of repetition at deeper structural levels is aired in some detail, and repetition underpins the symmetries within the *Ursatz* (see fig. 2.1). Reti (1951) was concerned with motivic connections, and believed that the unity of classical compositions (within and between movements) could either be ascribed to overt relationships between material, through: '*imitation*, [the] literal repetition of shapes, either directly or by inversion, reversion, and so forth, [or] *varying*, [the] changing of shapes in a slight, well traceable manner', or, more importantly, in his view, attributed to hidden associations, generated through '*transformation*, that is, creating essentially new shapes, though preserving the original substance, [or] *indirect affinity*, that is producing an affinity between independent shapes through contributory features' (p. 240).

Meyer's (1973) evolving reflections on musical patterning variously involve repetition, most overtly in his notion of 'conformant relationships', 'in which one (more or less) identifiable, discrete musical event is related to another such event by similarity' (p. 44). Although it is not stated openly, the concept is no less important, however, in the first chapter of *Music, the Arts, and Ideas* (1967), where the author's previously developed model of musical meaning is reviewed in the light of information theory. Moreover, Meyer identifies a number of different basic melodic structures (subsequently termed 'processes'—see Rosner & Meyer, 1986), including conjunct, disjunct and symmetrical patterns, whose internal regularity and use as stylistic archetypes imply repetition within and between works (Gjerdingen, 1988).

Repetition is also central to Forte's set-theoretical analysis, which, as we have seen, entails abstracting groups of pitch-classes and tracing similarities between them, and to semiological analysis, to which motivic similarities are fundamental at the paradigmatic stage. As Ruwet (1966/1987) says:

> I shall start from the empirical appreciation of the enormous role played in music, at all levels, by repetition, and I shall try to develop an idea proposed by Gilbert Rouget: '... certain fragments are repeated, others are not; it is on repetition—or absence of repetition—that our segmentation is based'. (p. 16)

Finally, consider that repetition ('parallelism') accounts for four of the five preference rules underlying Lerdahl and Jackendoff's (1983) 'generative theory of tonal music'. As the authors state: 'The importance of parallelism in musical structure cannot be overestimated. The more parallelism one can detect, the more internally coherent an analysis becomes, and the less independent information must be processed and retained in hearing and remembering a piece' (p. 52).

In relation to music analysis—the application of theory to a particular piece or group of pieces—Bent & Drabkin (1987) provide a useful summary:

> Analysis is the means of answering directly the question 'How does it work?'. Its central activity is comparison. By comparison it determines the structural elements and discovers the functions of those elements [...] comparison of unit with unit, whether within a sin-

gle work, or between two works, or between the work and an abstract 'model' [...] The central analytical act is thus the test for identity. (p. 5)

5. Hearing zygonic theory in action

So much for the recognition of repetition in music. To reiterate: zygonic theory takes a further step and suggests that it is a sense of *derivation* stemming from one musical element imitating another that is important in creating the sense of narrative in music. The easiest place to hear the theory in action is in musical 'canons', which are explicitly structured through repetition—one musical line consciously being made to copy another. Here, for example, is the opening of 'Et in unum Dominum' from the *B Minor Mass*, where Bach uses the derivation of the alto part from the soprano within a unified musical framework as a symbol of the Father *begetting* (not creating) the Son, which, according to Christian dogma (and Bach, by all accounts, was a devout believer), subsequently co-existed as parts of the same spiritual entity (fig. 5).

Irrespective of the symbolism, it is easy to appreciate how each note in the alto voice, ensuing shortly after an identical event sung by the soprano, sounds irresistibly to the musical ear as though it derives from it. In the mind, each pair of notes appears to be connected via a mental 'bridge' that spans the two perceived sounds. Each of these may be termed a 'zygonic relationship' or 'zygon' (after the Greek word for 'yoke', meaning the union of two similar things). In order to make analysis and understanding easier, it is sometimes helpful to represent these putative cognitive connections visually, and, at its simplest, this can be achieved through an arrow with a superimposed 'Z', as follows (fig. 6).

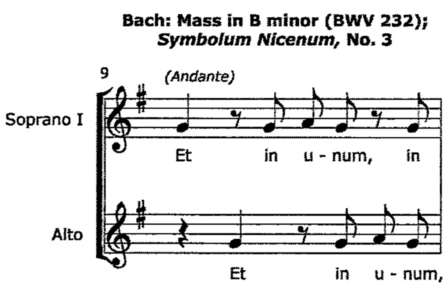

Figure 5 – Example of canon in unison.

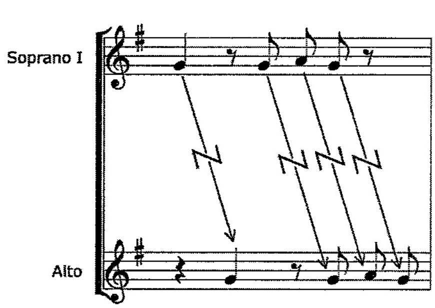

Figure 6 – Zygonic relationships.

Relationships *between* notes—for example melodic intervals—can also be repeated (fig. 7).

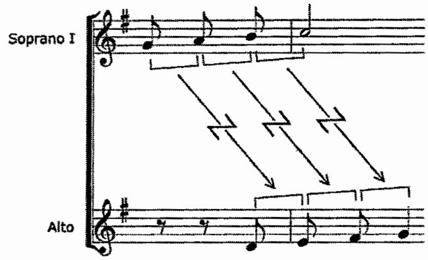

Figure 7 – Zygonic relationships between melodic intervals.

It is important to acknowledge that a zygonic relationship is merely a theoretical device—in Lakoff's (1987, p. 283) terms, a metaphorical representation of a particular 'link' schema—that inhabits the mental space pertaining to music processing (*cf.* Fauconnier, 1985/1994; Lakoff, 1987, p. 281 and 282), incurred by those listening to, performing or composing music. It is clear that zygonic relationships can at best offer only a highly simplified version of certain cognitive events that we can reasonably surmise take place during participation in musical activity. However, while simplification is always necessary to make headway in theoretical terms, some idea of the complexity involved can be gleaned by appreciating that the single concept of a zygon bequeaths a vast perceptual legacy, with many potential manifestations: between, for example, pitches, timbres, loudnesses, durations, interonset intervals (the lengths of time between notes), tonal regions, textures, processes and forms the same; over different periods of perceived time; and within the same and between different pieces, performances and hearings. There is, of course, no suggestion that the one concept is perceptually or even neurologically equivalent in all these manifestations, but *logically* so. Whatever the context, zygons may function in a number of ways: reactively, for example, in assessing the relationship between two extant values, or proactively, in ideating a value as an orderly continuation from one presented. They may operate between anticipated or remembered values, or even those that are wholly imagined, only ever existing in the mind.

Hence, empirical evidence in support of the theory is likely to be drawn from a diversity of sources. Currently, for example, one can point to experiments in auditory processing (such as the 'continuity illusion', summarised in Bregman [1990, p. 344 ff.]), and work on the perceptual restoration of omitted or obscured notes (see, for instance, DeWitt & Samuel, 1990), to support the presence of proactive zygonic-type processes. More recently, zygonic theory has been used to build a model of expectation in music (Ockelford, 2006), which has been tested experimentally with some success (Thorpe, Ockelford, & Aksentijevic, 2011; Ockelford & Sergeant, under review). And there is increasing evidence to support the theory in studies of children with learning difficulties, whose musicality tends to evolve in small steps, making the identification of the early stages in the development of music-structural cognition more straightforward than would otherwise be the case (Ockelford, 2008; Cheng, Ockelford, & Welch, 2009; Welch, Ockelford, Carter, Zimmermann, & Himonides, 2009; Ockelford et al., 2011).

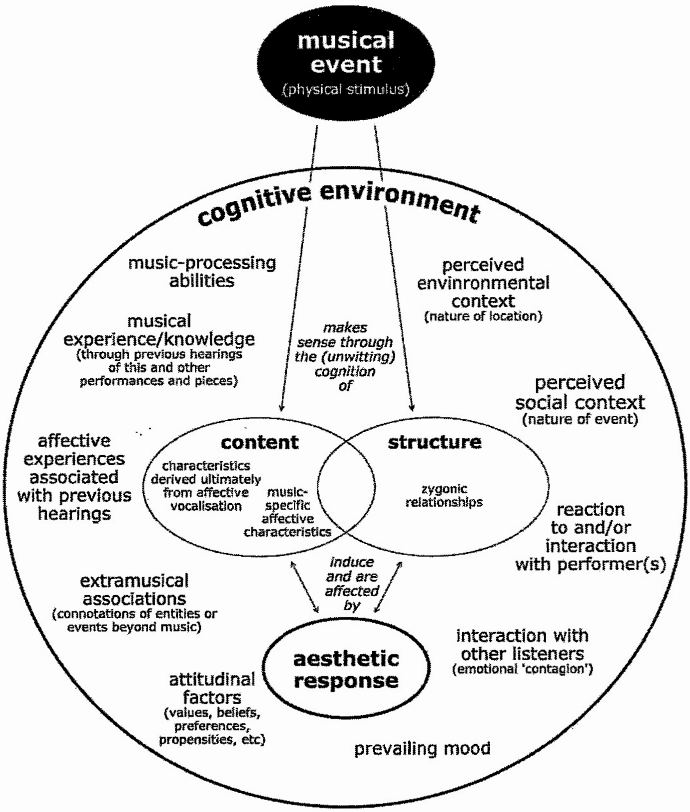

Figure 8 – The wider cognitive environment in which music-structural cognition resides.

Although, for the sake of theoretical simplicity, we have considered the creation and reconstruction of musical meaning in 'absolute' terms, in reality, these processes do not exist in isolation, but reside within and contribute to the 'cognitive environment' of listeners (Sperber & Wilson, 1995, p. 38 ff.). This is influenced to a greater or lesser extent by extramusical forces, pertaining both to the inner world of the person concerned (which is in turn determined by internal and external factors, past and present) and to his/her reaction to the immediate circumstances in which the performance is being heard. One extramusical factor that has been of concern to philosophers and psychologists alike, is the power of *association,* which can completely overwhelm a listener's reaction to intramusical attributes which, *ceteris paribus,* would occur, while nevertheless leaving intact his/her ability to recognise the sentiments which the piece would typically evoke in other listeners, and without compromising the internal 'sense' of the music. Hence, the wedding march played following the death of one's partner may still be recognised as essentially joyful (even though it may elicit intense grief) and be perceived as musically coherent (even though its effect in aesthetic terms is the opposite of that which the composer intended).

Other factors pertaining to listeners include the emotional and aesthetic range of experiences they bring to bear; their knowledge of music, gained through previous hearings of this and other performances of the current piece and others; 'extramusical associations' (connotations of non-musical entities or events established through previous experience that may be stimulated by further hearings of a piece or feature of it); their music-processing abilities; attitudinal issues, such as values, beliefs, preferences, and propensities; and their prevailing mood, which will provide the affective backdrop against which any emotions aroused by the music will be superimposed as phasic perturbations (Davidson, 1994). The external environment can influence aesthetic response in a number of ways too. A listener may well be affected by the behaviour of the performer and by the reactions of other people who are present, through empathy and 'emotional contagion' (Scherer & Zentner, 2001, p. 370). Other considerations include the social context in which the music is being heard and the nature of its location (*ibid.,* p. 364-365). All these factors contribute to the cognitive environment of the listener, which may be represented schematically as follows (fig. 8). This shows the central place that the cognition of structure plays in the listening process as a whole.

6. Example of zygonic theory in action—the first movement of Mozart's piano sonata, K. 333

Zygonic theory suggests that *every* aspect of music is supersaturated with repetition in a way that goes far beyond the motivic and thematic replications and variations that traditional music analysis seeks to identify. Take, for example, the first movement of Mozart's Piano Sonata in B^{b}, K. 333 (*cf.* Ockelford, 2005, p. 35-66). To the ear familiar with Western classical music, nothing could sound more natural, more unpretentious—and less like a multidimensional matrix of meticulously crafted logical connections in sound. Yet zygonic analysis indicates that there are as many as 40 types of structural imitation in this movement potentially operating at any given time (Ockelford, 1999, p. 704 ff.), connecting individual pitches, the intervals between them, and their relative probabilities of occurrence; harmonies, harmonic transitions, forms of dissonance, and the manner in which these are resolved; keys, modulations, and longer-term tonal structures; the durations of notes, their interonset intervals, metre, rhythms, phrase-lengths, and lengths of sections; textures, melodies, the connections between them, and overall form. Compromise the structure in any one of these dimensions and the musical fabric as a whole would no longer hang together coherently. Of course, that is not to say that Mozart gave any of this a moment's conscious thought as he composed K. 333. For him, by all accounts, creating music was as natural and effortless as speaking. But just as producing and understanding cogent and expressive verbal language makes huge demands on cognitive processing (of which speakers and listeners are typically wholly unaware), so too does generating and attending to music.

We will consider first the 'background' organisation of the movement: general characteristics that are more or less common to other pieces in a similar style. Hence they do not determine what makes a piece unique (it is largely the nature of relationships between successive events that give a work its exclusive identity), but rather facilitate memory, frame understanding and fuel expectation (Huron, 2008). It is in the domains of pitch and perceived time that the great majority of background structure, and by far the most musically sophisticated and significant, occurs (Ockelford, 2005, p. 37). For example, 78% of interonset intervals between successive notes are in the ratio 1 : 1, implying vast networks of potential zygonic relationships (fig. 9). (Further work would be required to determine which of these would be likely to be realised in cognition.)

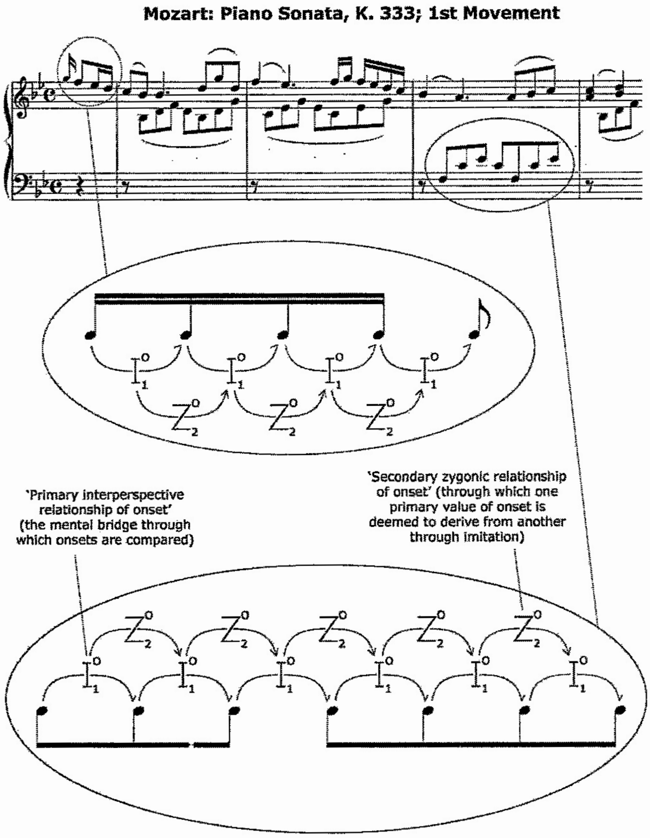

Figure 9 – Examples of the potential zygonic relationships linking interonset intervals the same in the first movement of Mozart's Piano Sonata, K. 333.

Other interonset ratios are encountered far less frequently, as the chart pertaining to the first movement of K. 333 shown in Figure 10 shows. Similar distributions are found in other piano sonatas by Mozart, implying the existence of perceived temporal imitation between pieces operating on the statistical level. In K. 284, K. 310, K. 311 and K. 333 the degree of similarity in the distribution of interonset ratios is 92% (where dissimilarity is calculated as the average divergence from the mean of each category of ratio):

$$\text{Similarity (\%)} = 100 - \sum \left(\frac{\sum \left| x_i - \left(\frac{\sum x_i}{n}\right) \right|}{n} \right)$$

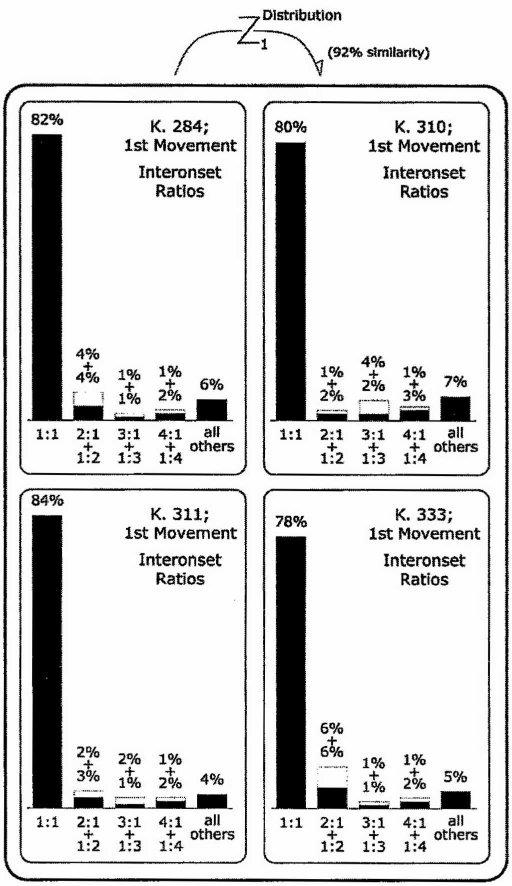

Figure 10 – Imitation of interonset ratios within and between the first movements of piano sonatas by Mozart.

Similarly, in the domain of pitch, an analysis of melodic intervals between adjacent notes reveals substantial background organisation across the first movement of K. 333 as a whole. A little over 90% of intervals are a perfect 4th (five semitones) or smaller. Of these, major 2nds (two semitones) alone account for almost 40% of all melodic transitions (fig. 11). Moreover, the essential characteristics of this distribution are a feature of other pieces too, implying stylistic imitation of the type shown. Moreover, the tendency of small intervals to occur much more frequently than large ones is by no means confined to the music of Mozart, as a number of studies pertaining to various Western genres have shown. These range from folksongs to many styles of classical music and popular music of the twentieth century (see, for example, Fucks, 1962; Jeffries, 1974, p. 904; Dowling, 1978, p. 351 and 352; Huron, 2008, p. 74 and p. 158-161).

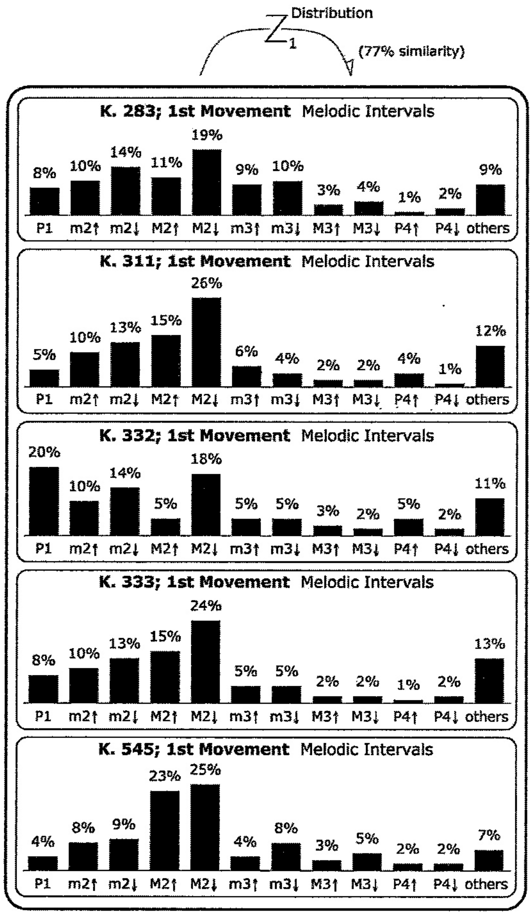

Figure 11 – Imitation of melodic intervals within and between the first movements of piano sonatas by Mozart.

Beyond this, zygonic analysis shows how the logic of the music and its expressive character—the 'carrier' and the 'message'—are fused in an abstract discourse that is wholly persuasive. Consider, for example, Mozart's use of appoggiaturas in the opening phrase. Zygonic theory contends that the fourth appoggiatura, heard structurally as a product of the first three, does more than *replace* their affective qualities in current consciousness: it *transforms* them. That is, the relative stability of the opening descent (comprising three pairs of notes that unfold the major tonic triad) is remodelled to generate the conclusion to the second phrase, consisting of a further appoggiatura that resolves onto the supertonic minor harmony—a metamorphosis that engenders a sense of yearning that the minor chord in isolation would not have produced (fig. 12).

Intuitively, to resolve the miniature aesthetic narrative built up to this point, the appoggiatura needs to be transformed again, and made to re-appear in the context of a tonic harmony. Mozart achieves this through compelling though unobtrusive structural logic (fig. 13).

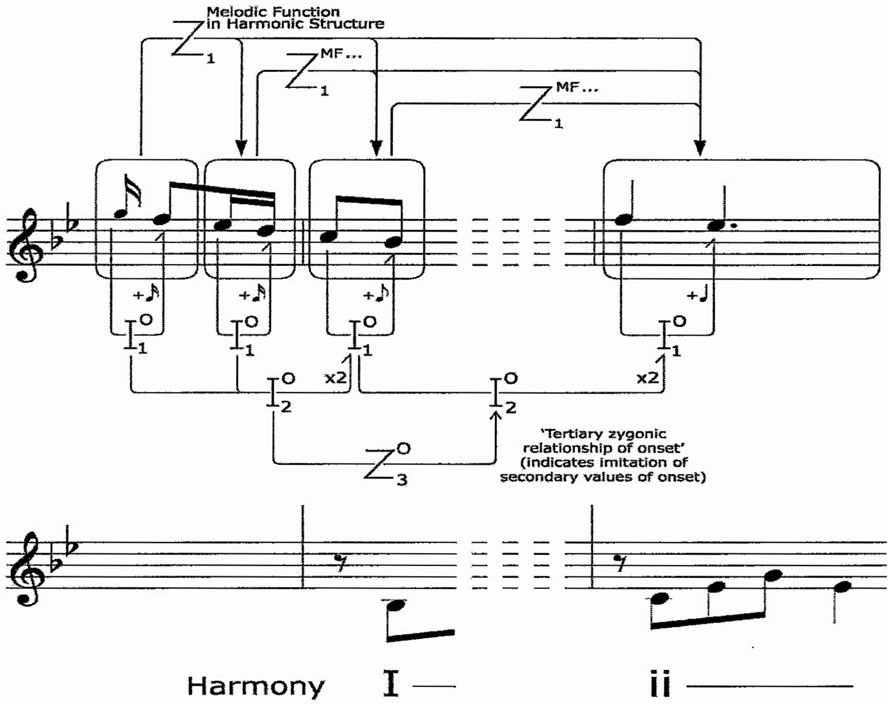

Figure 12 – The structural and affective use of appoggiaturas in the first two bars of K. 333.

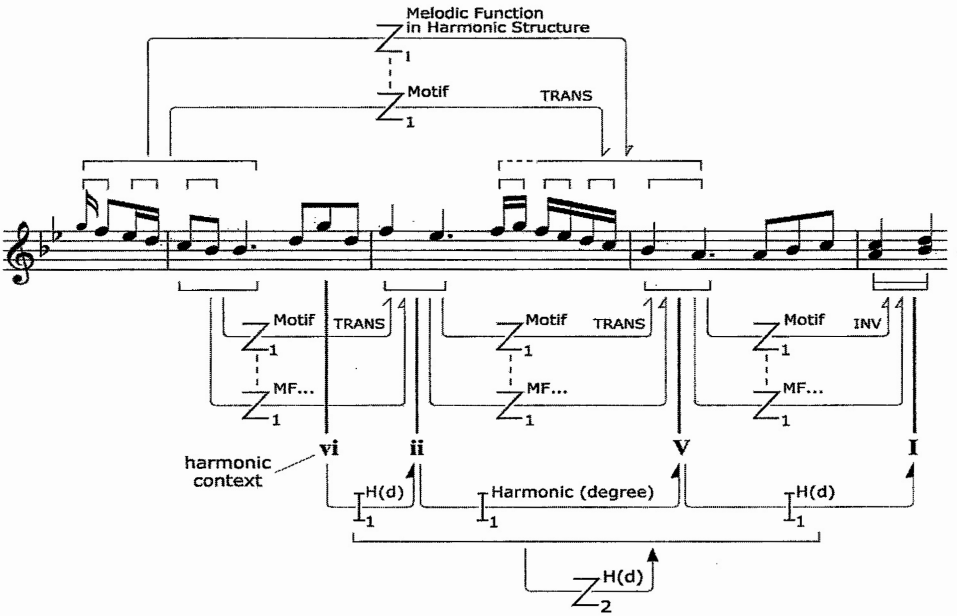

Figure 13 – Systematic relationships between appoggiaturas and harmonic and melodic structure in the opening bars of K. 333.

The opening phrases are linked in other ways too, simultaneously implying retrogression on one level and transposition on another (fig. 14).

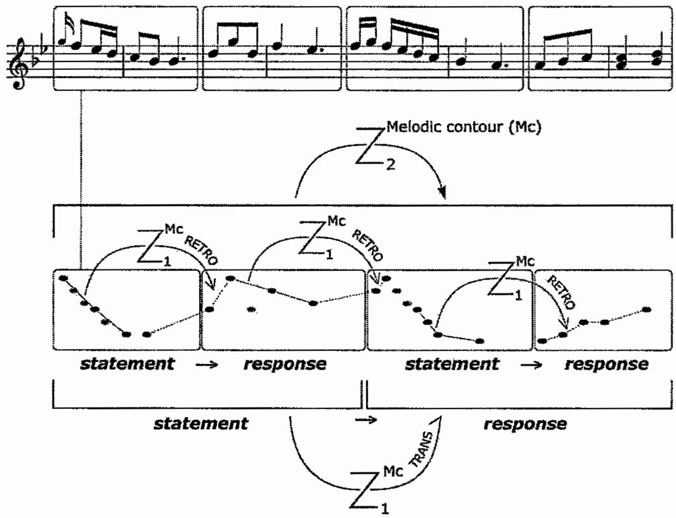

Figure 14 – Effect of statement and response at different levels in the structural hierarchy accomplished through complementary forms of zygonic organisation.

The tight though understated structure that characterises the first subject of K. 333 continues as the movement unfolds, not only within sections but between them. For example, the second thematic group (from bar 23) is closely related to the first, both in its descent from the sixth degree of the scale and its identical use of appoggiaturas (fig. 15).

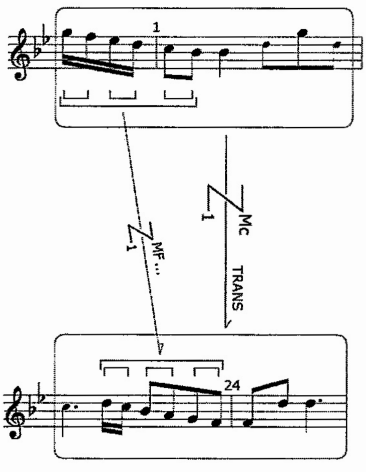

Figure 15 – Relationship between the opening themes of the first and second thematic groups in the first movement of K. 333.

Then, the pitch range of the opening melody (from leading note up to submediant) is respected on a number of subsequent occasions, for example from bar 15. In bar 43, the interval explicitly becomes part of proceedings, forming the climax of the passage that begins at bar 39 (fig. 16). Hence, once more, we find evidence that the expressive content of the music is held together through zygonic relationships functioning in addition to those demanded by the structure of the classical sonata, imbuing the music with a sense of aesthetic unity that formal connections alone could not guarantee.

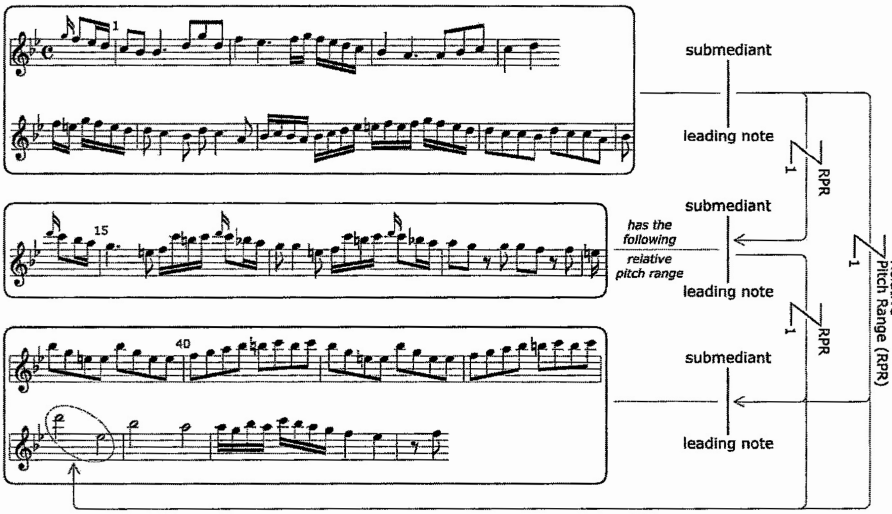

Figure 16 – Use of melodic range as a unifying feature in the first movement of K. 333.

7. From theory to metatheory—what makes music 'music'?

It may be thought that the pervasive use of imitative structures like those like found in K. 333 is exceptional, but the analysis of a wide range of music—from folksongs to fugues—across many different cultures (see, for example, Ockelford, 1999) suggests that it is actually typical. Conversely, it is impossible to imagine a scenario in which 'music' was to be created in which there was no sense of contingency between its constituent sounds. There would be

> nothing but a haphazard succession of irregularly varying pitches which, differing widely in duration and time of attack, would describe wildly irregular rhythms, unpredictable in loudness, coloured by a bewildering confusion of timbres and proceeding indiscriminately from random locations potentially spaced the world apart. (Ockelford, 1999, p. 1)

Hence it seems reasonable to hypothesise that imitative structure may be a universal feature of music. One way of testing this argument is to consider it alongside other definitions of music, then, in the light of these findings, to use zygonic theory to interrogate the notion of music itself, and finally to ascertain how the proposition fares in relation to potentially exceptional cases.

The question of 'what is music?' has been regarded a legitimate area of philosophical concern since the time of the ancient Greeks, and there are those that contend that, since music has varied (and continues to vary) so much with time, place and culture, it would be impossible to arrive at a single, fully-inclusive definition of music (*cf.* Molino, 1975, p. 37). As Nattiez (1990) says: 'By all accounts there is no single and intercultural universal concept defining what music might be' (p. 55). However, that is not the same as arguing for a more limited form of universality, whereby there may be *aspects* of music that are necessary to its existence, even if they are not sufficient to offer a comprehensive definition in all contexts. If so, what might some of music's universal features be?

One contender must surely be Varèse's all-embracing notion of 'organised sound', which he coined in the 1920s in an attempt to broaden the generally-accepted concept of 'music' that was prevalent in the West at the time, to include his own experimental work. The controversy this caused had less to do with the idea of musical elements needing to be 'organised' than Varèse's inclusive definition of 'sound' in a musical context, which encompassed what many people took to be 'noise': that is, sounds they did not like and which therefore could not, in their view, constitute 'music'—the art of the beautiful (see Varèse & Wen-Chung, 1966, p. 18).

Almost one hundred years later on, the universe of sounds that people regard as potentially 'musical' has widened considerably, and it is easier to focus on the real problem with Varèse's definition: that organised sound does not have to be 'musical'. In his ecological approach to auditory event perception, Gaver (1993) points out that there are essentially two ways in which we hear sounds: through 'musical listening', which focuses on perceptual qualities such as pitch and loudness, and via 'everyday listening', which is more concerned with function. To these two categories, we should add the organisation of sound as language (Scruton, 2009, p. 4). Hence Varèse's definition of 'organised sound' can be seen to suffer from 'overextension', in that it does not apply exclusively to music.

Scruton's (2009) account takes a further step, by making it clear that music "relies neither on linguistic order nor on physical context, but on organization *that can be perceived in sound itself, without reference to context or to semantic conventions*" (p. 5, italics added). But by resolving one issue, Scruton raises another: that of 'perception'. Whose perception? As Lerdahl (1992) points out, the organisation that composers utilise in their work need not be detectable by listeners. Moreover, because of the high degree of redundancy that typically characterises the musical message (Moles, 1968; Cohen, 1962), there may well be structures that can be identified in pieces that were incidental to the process of composition (Ockelford, 2009b, p. 86). Indeed, that special breed of listener—the music analyst—is capable of identifying structures that lie beyond practical levels of engagement with music (Ockelford, 2009b, p. 88).

Thus what is music to one person's ears need not be music to another's. Indeed, it is conceivable that material that was never intended to be construed as music (ranging from the sounds of nature to the noise of machines) could be heard as such. Hence we appear to have argued our way back to the relativist position that what constitutes music is purely in the ear of the beholder, and that so-called 'universals' are necessarily illusory constructs pertaining to the prevailing culture of discourse. But this feels counter-intuitive. The very fact that it is possible to make observations such as the one with which this paper opened (that the great majority of people apparently engage with music every day of their lives) implies that there is a single concept, or, at least, a bundle of concepts, to which most of us can subscribe. So let us see if we can escape the circularity of the arguments that have been advanced by analysing what makes music 'music' using zygonic theory. Such thinking runs as follows.

Zygonic theory holds that:

- 1. the essence of music is that one sound or group of sounds, or feature or features thereof, should be heard as deriving from another or others through imitation. Hence music is a purely cognitive phenomenon, and while there may (but need not) be physical correlates of our internal audition, these do not constitute 'music';

- 2. a sense of derivation through imitation enables us to hear discrete sonic events as a coherent stream of abstract sounds, as 'music': and just as each event has the capacity to induce an emotional reaction, so the contingencies we hear in a series of musical sounds can evoke an emotional narrative that unfolds in time. (This is distinct from musical meanings that derive through association.);

- 3. almost without exception, mature humans have the capacity to hear sounds and the relationships between them as being derived from one another through imitation; this requires no formal education, and typically occurs non-consciously: we are virtually all intrinsically 'musical';

- 4. a sense of derivation through imitation is a necessary feature of all structures that we perceive as musical; a consequence of this is that all music is infused with repetition in all domains and at all levels;

- 5. hence music is typically supersaturated with far more repetition than is required for it to be coherent, and this has two consequences: (a) listeners do not need to hear all the available structure for a given musical message to make sense, and (b) different listeners (or even the same listener on different occasions) can apprehend different structural elements, yet each can still have a coherent musical experience. Moreover, it is possible that a sense of derivation through imitation that was conceived by a composer may not be detected by listeners, and *vice versa*. Nonetheless, there is normally enough common perceptual ground for pieces of music to exist as shared and meaningful cognitive enterprises;

- 6. music can and may well be associated with other social and communicative activity (such as dance and verbal language), which may interact with the cognition of purely musical structures and contribute additional layers of meaning. In some cultures, the notion of 'music' embraces more than just abstract patterns of sound. Nonetheless, streams of sound, structured through a sense of derivation through imitation, are a feature of all the musical experience.

We will now test this definition against a number of different scenarios.

Take, for example, the notion of music in nature (for example, Rothenberg, 2001; Sallis, 2006, p. 56) and consider (for instance) Tennyson's 'babbling brook'. This may be regarded by some as 'music to the poet's ears', but zygonic theory would contend that it can be so only in a metaphorical sense, since there is no imitation present; no coherent narrative in which sounds are deemed to derive from one another through imitation. Hence rushing water (or similar features of a naturally occurring soundscape, such as pattering rain and wind sighing through the trees) fail the 'musical' test (fig. 17).

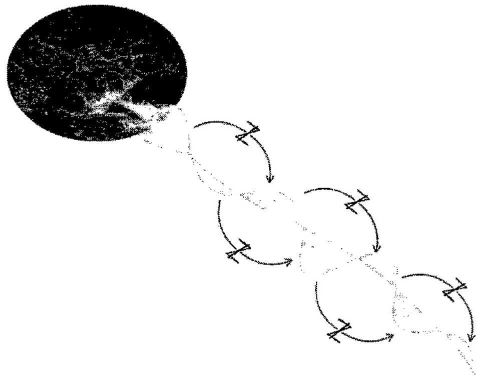

Figure 17 – The sound of the babbling brook does not create patterns that can be considered to be derived through imitation.

With another natural sound—birdsong, the position is more complex. Birds typically use short, distinct motifs that are well defined in terms of pitch, time and timbre, and which as youngsters they have usually learnt from a conspecific adult (Phan, Pytte, & Vicario, 2006; Gobes & Bolhuis, 2007). They often repeat their calls to form chains of avian melody. But do such concatenations constitute music? Consider, for example, the following series of four 'cuckoo' calls recorded in a natural setting (fig. 18).

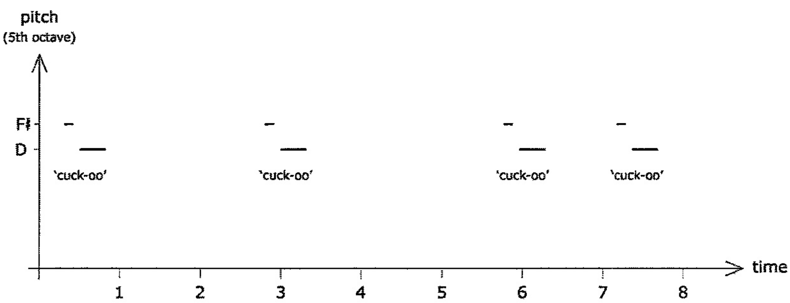

Figure 18 – Representation of four 'cuckoo' calls recorded in a natural setting—do they constitute music?

It is, of course, impossible to say for sure whether the bird acted with any sense of self-imitation as he sang to defend his territory, or whether the repetition that occurred was merely a by-product of his communicative instincts. There is evidence that cuckoos have some awareness of their vocal products, since they can distinguish their own songs from those of other species. However, since the development of their singing need not have been influenced by conspecific adults, and as their calls are very similar across a wide geographic range, it appears that they are innate rather than learned (Payne & Sorenson, 2005, p. 97), and therefore *not* deliberately imitative. However, it is perfectly conceivable that humans may hear the successive 'cuckoos' illustrated in figure 20 as being derived through imitation, implying that the series of calls constitutes, in their minds, music (fig. 19), in effect constituting a reverse example of a Lerdahl's (1992) mistmatched 'compositional' and 'listening' grammars.

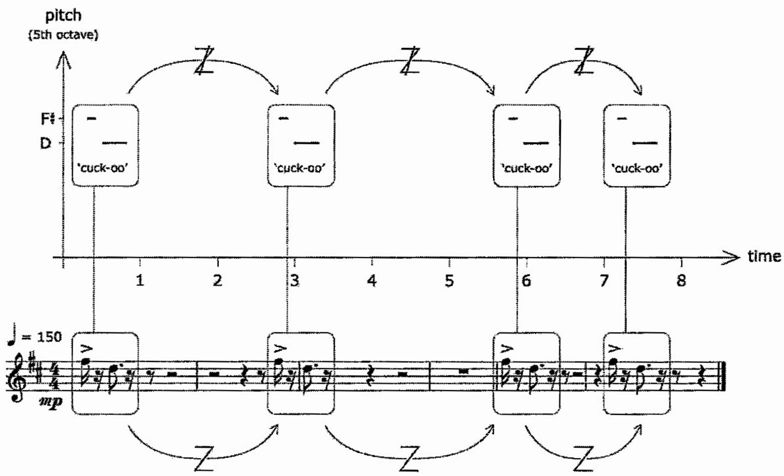

Figure 19 – A listener hearing cuckoo calls as music, despite derivation through imitation not being part of the bird's thinking.

So is the sequence music or not? Referring to Clauses 1 and 5 above, the repetitive structure of the succession of calls means that they *have the capacity to be heard as music*, even though they were produced (we assume) with other-than-musical intent. Hence, the classification of the sequence as music will depend on the knowledge and beliefs of the listener. If this appears to be unnecessarily open-ended, a further scenario will show that a certain conceptual fuzziness is inevitable. Consider another set of four 'cuckoos', similar to those shown in figure 20, but a major third lower, more equal in duration, and with shorter time periods between each pair of sounds. These differences should not in themselves produce phenomenological change of any substance, so the ontological arguments set out in relation to the sounds' existence or non-existence as music still apply. But now consider that the tone-colour is that of a B♭ clarinet. For those familiar with the timbre, it is unequivocally indicative of a musical instrument and therefore, very likely, a sense of human agency. The chances of the motifs being heard as deriving through imitation—of the passage being heard as music—increase markedly. Finally, imagine a context in which the cuckoo calls are heard in combination with a nightingale's song (played on the flute) and a quail's (on the oboe), and appear in the context of the slow movement of a classical symphony (fig. 20). To the culturally attuned ear, the implication of imitation is now irresistible, and representations of the sounds of nature are definitively 'music'.

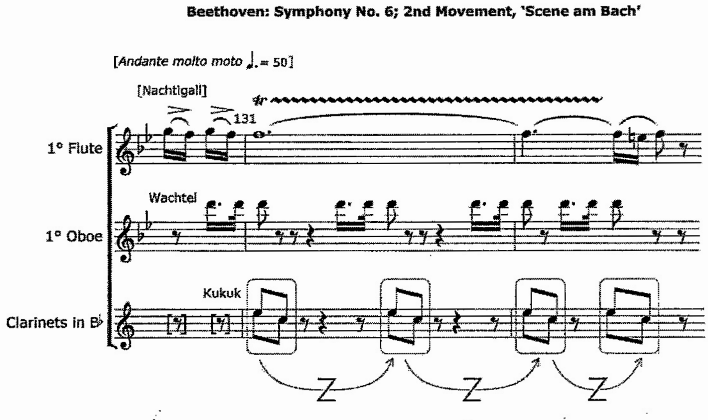

Figure 20 – Cuckoo calls used in the context of a symphony ensure they are heard as music.

We now move on to consider music and language. As figure 3 shows, language generally fails the 'musical' test, since it is driven by semantics rather than the imitation of the sounding qualities of words. However, exceptions do exist: rhyme, assonance and alliteration, for instance, provide examples where musical logic has arguably encroached upon the realm of verbal language. The regularity of metre in poetry can also be interpreted as imitation in sound and therefore construed musically. Whether or not listeners *do* hear phonetic and metric repetition in a musical way remains to be determined. Some indirect evidence does exist, however. For example, it seems that elements of linguistic and musical syntactic processing may share neural resources (Patel, 2008, p. 285), and empirical work has suggested that in songs, music and language may be encoded together (Serafine, Crowder, & Repp, 1984; Morrongiello & Roes, 1990), although expert singers may have the capacity to decouple the two forms of auditory communication in neurological terms (Wilson, Abbott, Lusher, Gentle, & Jackson, 2010). Then, the presence of a melody can increase phonetic recognition (Lebedeva & Kuhl, 2010), and a tune can facilitate the learning and recall of attendant words, *provided that the music repeats* (Wallace, 1994, p. 1471).

Given this, if aspects of language alone were being processed musically, we could reasonably hypothesise that verbal strings that were structured quasi-musically (for example, through the imitation of words' sounding qualities, as in poetry) would be learnt and recalled more easily than in the inherently 'non-musical' prose. And there is indeed evidence that poetic forms support memory (Rubin, 1995). Furthermore, it appears that the cognitive advantages of 'word-music' may not be confined to learning and recall: alliteration has been shown to aid verbal comprehension (Lea, Rapp Rapp, Elfenbein, Mitchel, & Romine, 2008), for example, and the sounding features of words appear to influence lexical selection (Rapp & Samuel, 2002).

Other, more direct, evidence of language being processed musically can be found in the world of children whose development occurs in the context of what I term 'exceptional early cognitive environments', incurred through severe visual impairment or autism (Ockelford, 2011). Given the social and communication challenges that both blindness and autism pose, it is perhaps unsurprising that in the case of a significant proportion of children with either of these disabilities, language acquisition can be delayed or aberrant (Fay, 1973). A characteristic of speech frequently found in both groups is so-called 'immediate echolalia'—the prompt repetition of words or phrases that often has no obvious semantic or pragmatic function (McEvoy, Loveland, & Landry, 1988; Mills, 1993, p. 163). Elsewhere (Miller & Ockelford, 2005), I have suggested that echolalia may arise in part as a consequence of words being treated as musical objects, whereby they are structured primarily through imitation, rather than being treated syntactically as language.

This has correlates with other behaviours seen in some children on the autism spectrum with little or no language, which include treating environmental objects as potential music makers and everyday sounds as music (Ockelford, 2010a). For example, Freddie, aged 10, enjoys emptying out flowerpots from the patio, arranging them rather like an earthenware gamelan in the kitchen, and then playing them with rapid finger-flicks. Meanwhile, Romy, aged 8, incorporates the sounds of aeroplanes and mobile phone ring-tones into the pieces she plays by ear on the piano—rather like latter-day Beethovenian cuckoos (Ockelford, 2010b).

Parents report that some young blind children too can become very interested—even obsessed—with certain sounds in the domestic environment, such as the humming of the microwave or the whirr of the tumble-drier (Ockelford, Pring, Welch, & Treffert, 2006). Moreover, the accounts of parents of children with no sight suggest significantly higher levels of early interest in music than in those who are partially or fully sighted (Ockelford & Matawa, 2009; Ockelford, Gott, Risdon, & Zimmermann, 2011). Finally, the prevalence of absolute pitch (AP) in blind children (Pring & Ockelford 2005; Ockelford, 2008; Dimatati, 2009; Dimatati, Ockelford, Heaton, Pring, & Downing, 2010) and autistic children (Heaton, Hermelin Hermelin, & Pring, 1998; Heaton, 2009) arguably indicates a cognitive preference for the *musical* processing of sounds. In this regard, it is of interest to note that AP frequently manifests itself around the age of 24-30 months (Ockelford, 1988) – the point at which general development, including the evolving use of language in children who are autistic or blind, so often seems to become derailed (Lösche, 1990; Dale & Salt, 2007, p. 685), reinforcing the notion that 'early atypical processing biases may play an important role in the development of splinter skills' (Heaton, Williams, Cummins, & Happé, 2008, p. 216).

Turning now to Morse—a specialised form of encoded language—the poetic niceties of rhyme, assonance, alliteration and metre are, of course, eliminated, and with them, any phonetic and metric similarities that could potentially have been heard in a musical way. However, as the system uses only one (steady state) pitch and two different durations, there is an immense amount of repetition. Could this be construed as music? To those using Morse code as a form of communication, the answer is usually 'no', since any replication is driven either by the design of the code (for example, the three dots or 'dits' that make up the letter 'S' and the three dashes or 'dahs' that make up the letter 'O'), or by semantics (as in the repeated 'S'—'dit-dit-dit'—in 'SOS', for instance) (fig. 21).

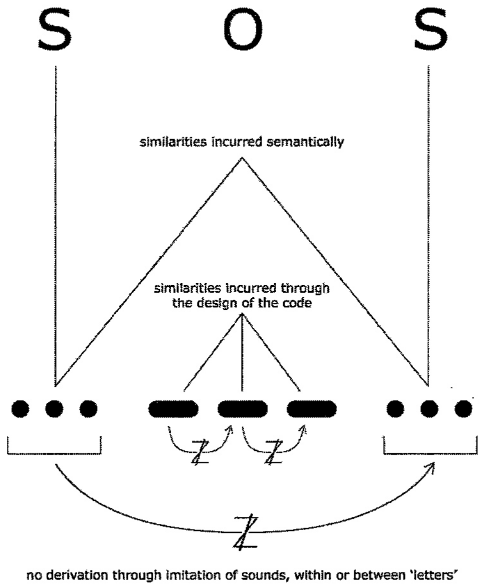

Figure 21 – Morse code heard purely as a means of communication does not constitute music.

For those who do *not* recognise the code as a representation of verbal language, it is conceivable that it may be interpreted in a musical way, though the irregular additive nature of the temporal structure (in which the duration of a 'dah' is 3 'dits', the interval between sounds that make up the same letter is equivalent to one 'dit', and the time *between* letters is three 'dits') makes it difficult to parse events other than at the surface level, and the rapid speed at which individual sounds pass by militates against this.

However, the fact that Morse uses materials that resemble the notes that are the building blocks of much music means that it has been relatively easy for composers to incorporate the code—or, at least, readily recognisable transformations of it—into pieces. Such modifications enable the syntactic demands of music for repetition and the semantic requirements of Morse to correspond to a series of letters that make up a word to be reconciled. Take, for example, the theme from a UK television series (1987-2000) featuring the enigmatic Oxfordshire detective Inspector Morse. The composer, Barrington Pheloung, captured something of the enigma of the central character by incorporating his name into the music as follows (fig. 22). This was achieved by playing the code around a quarter speed, changing some durations to permit metrical regularity, and sounding the entire M-O-R-S-E sequence repeatedly as an ostinato. These strategies mean that, while enough of the identity of code was retained for it to be interpreted semantically, it could also be heard as pure sound, structured through imitation that occurred both within the M-O-R-S-E sequence and between its many appearances.

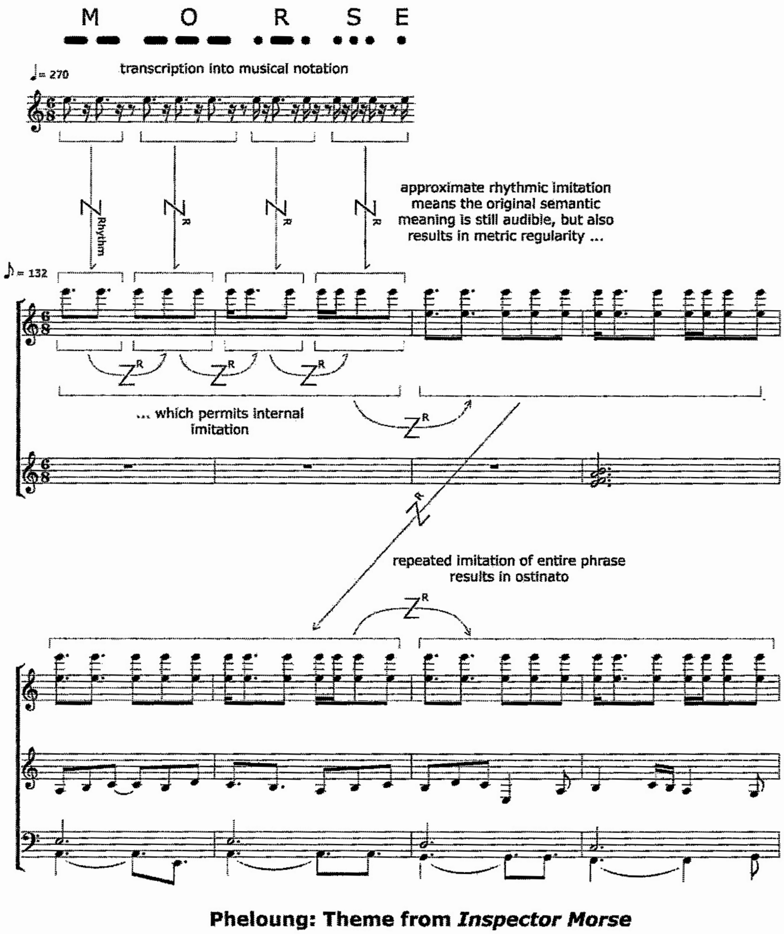

Figure 22 – Incorporation of Morse code into music results in 'conceptual blending'.

Hence the full meaning arises from a 'conceptual blend' (Fauconnier, 1985/1994; Fauconnier & Turner, 2002; Zbikowski, 1998, 2002; Cook, 2001), created by the composer, and which can only be appreciated by those listeners who are able to process the structures of both the Morse code and the music. A reverse example, of a conceptual blend that was *not* planned by the composer, but that was subsequently foisted upon the music by listeners, occurs in the opening four notes of Beethoven's fifth symphony. These correspond to 'V' in Morse code ('di-di-dit-dah'), and, in 1941, the BBC started to use the opening of the symphony as a theme for radio shows beamed across Europe, in the hope of reminding people of Winston Churchill's famous two-fingered salute, and so boost morale during the Second World War. Here, what was originally pure music, subsequently acquired a semantic overlay.

8. Conclusion

This paper has shown how zygonic theory offers an explanation of how music makes sense and how it conveys meaning in the absence of semantics. The zygonic model which predicts that repetition should be universally pervasive in music appears to be confirmed. It provides the means of addressing some philosophical issues as to what can reasonably be considered to constitute music and what cannot. It does not preclude the notion of music as defined forming one strand in wider social and communicative activity such as song and dance.

While some empirical evidence exists for the claims that are made, further research is needed to test the assertions that are set out. For example, zygonic analysis needs to be undertaken of music from a wide range of cultures, and the results discussed and compared with ontological conceptions held by its creators and original audiences. There needs to be work on how to distinguish between the recognition of repetition and the sense of derivation through imitation at the heart of zygonic theory. And there needs to be investigation—including, if possible, neuroscientific studies—of whether some children with autism or blindness or both do indeed process language and everyday sounds as music, with a view to devising interventions that may reduce the risk of early development becoming derailed, as well as seeking to shed further light on music cognition among 'neurotypical' populations.