

MIDDLE-OUT MUSIC ANALYSIS AND ITS PSYCHOLOGICAL BASIS

Richard Parncutt¹ and Robert Pascall²

¹Department of Music, University of Graz

²Department of Music, University of Wales, Bangor

ABSTRACT

Nattiez advocated bottom-up analysis, in contrast to Ruwet's top-down; Schoenberg's seminal work on *motive* engaged with the musical surface; and Schenker did not differentiate between hierarchical voice-leading levels as regards perceptual salience. A critical appraisal and comparison of these contrasting music-theoretic approaches suggests that they may all be consistent with an analysis that begins from the most salient hierarchical level – the phrase and equivalent unitary gestures or figures, which we call *the middle level* – and moves outward to faster and slower levels. We suppose that the main cognitive processes active at, and characteristic of, the middle level are *parsing* (segmentation and functional interpretation of segments/chunks), and triggering or deployment of style schemata (*topics*). We review literature that supports the psychological reality and salience of the middle level, showing how it co-ordinates and focuses analytic activity around a fundamental human cognitive modality, ensures hearing relevance in the analytical outcome, and relativizes divergent analytical approaches.

1. INTRODUCTION

The incorporation of music psychological concepts into music analysis, at least at an intuitive level, can focus and legitimate music-analytical techniques. Meyer (1956, 1967, 1973), Narmour (1977), and Lerdahl and Jackendoff (1983) are among the music theorists to have developed links between music theory and music psychology, and to have explored their music-analytic applications.

Here, we explore the psychological basis for a listening and composing strategy that privileges a middle level of perception: neither too detailed (melodic and rhythmic figures) nor too long-term (formal archetypes). We first enquire about the relative salience of different hierarchical levels of analysis, from the fastest (surface) to the slowest (large-scale form). Convergent psychological evidence suggests that analytic levels vary in their perceptual importance, so a cognitively relevant analysis should first concentrate on the most perceptually relevant level or levels.

We are not necessarily talking about how a piece is experienced in a first hearing, or even after several hearings by an expert listener. Instead, we are talking about how analytic knowledge can enrich the hearing of a musical work, rather than misleading the ear by appealing to principles that lack a clear psychological foundation. We predict that this approach will yield optimally useful analytic strategies.

2. MUSIC-THEORETIC PRECEDENTS

Schenker did not regard the fore-, middle- and background levels as unequal in their perceptual and compositional relevance or salience. His successive *composing out* of layers from the slowest to the fastest reflects a concern with elaboration as a structuring device, without predetermined perceptual preferences between hierarchical levels. While his focus of attention was mostly on voice-leading to the relative exclusion of thematic concerns, he did, in the latter stages of *Der freie Satz* (1935/1979), increasingly concern himself with the relation of voice-leading analysis to the theory of form, and began to re-legitimize the roles of thematic contrast and development as form-building elements.

Other analysts more explicitly acknowledged differences in salience among analytic levels. Schoenberg's (1943) seminal work on *motive* engaged primarily with the musical surface. The motive consisted of "one or more features of interval and rhythm" and was understood to be a generator and guarantor of coherence operating across a whole piece. He maintained that the motive appears unequivocally at the beginning of a piece and recurs throughout in differing types of variation. It therefore acts to bind basic segmentations, but cannot itself provide that segmentation (cf. Cambouropoulos & Widmer, 2000).

Schoenberg's approach may be regarded as a form of *phrase analysis* – a topic with which theorists have been concerned since at least Koch (1782-93). Phrase analysis was central to Schoenberg's pedagogy, but marginal in late Schenker, whose thinking was primarily directed towards laying bare the counterpoint of harmony through hierarchical voice-leading structures, using prolongation and goal-motions. In marginalizing the constructive roles of theme and motive, he did not, however, negate the power of phrase, as his analysis of the C-major Prelude from Book 1 of Bach's *Well-Tempered Clavier* (Drabkin, 1985) illustrates. Again, the most salient level of analysis lies not at, but between, the surface and the background. It appears that the approaches of Schenker and Schoenberg to musical structure, although quite distinctive, can be realistically integrated into a pan-theoretic stance. If phrase forms the emblematic component of the middle level, we can co-ordinate Schoenberg's concern with *motive* (how phrases are connected with one another by means of internal and small-scale similarities and variations) and Schenker's concern with *voice-leading* (how harmonic prolongation and goal-motion at every level is both elaborative and form-building) as ways of understanding the sequencing of middle-level units.

The relationship between hierarchical levels is a common theme in music analysis. Ruwet's (1966) analyses began with the whole piece and moved progressively to smaller and smaller segments. His initial segmentation was determined by the largest-scale repetition within the piece. He assumed that repetitions within a melody are structurally equivalent to the recurrence of words within and between sentences in natural language. His method worked best for short pieces and has yet to be applied to pieces of the scale of classical sonata movements. Nattiez, in his extended analysis of Varese's "Density 21.5" (1982), called for the "most exhaustive inventory" possible of all recognizable units within a piece, and began his own analysis with the earliest small-scale repetition, to build up a multi-layered account. He explicitly contrasted his "bottom-up" method with Ruwet's "top-down". His "esthetic" analysis does not, however, concern itself with psychologically-based strategies for listening.

3. THE MIDDLE LEVEL

Clearly, music theorists often intuitively – if not explicitly – regard the phrase as the focus of the listener's attention, and thus as the most salient perceptual level of a composition. Here, we wish to raise this idea to the level of an analytic paradigm that is susceptible to scientific-psychological testing. We specifically propose that there exists in most music a middle level that is more perceptually salient or relevant than all faster and slower levels. We further propose that a cognitively relevant analysis should begin from this level and move outward – a *middle-out analysis*.

Music-psychological research on rhythm frequently refers to the *tactus*, or most salient pulse sensation evoked by a piece of music, corresponding to a moderate tempo (Lerdahl & Jackendoff, 1983; Parncutt, 1994). Our middle level is *not* a *tactus*, nor is it a metrical level (although there is certainly an interaction between metrical levels and the middle level; Rothstein, 1989). We regard pulse and metre as a gestalt-psychological background that is established at the start of the piece and remains much the same throughout – at least, for the kind of music with which we are primarily interested. The listener's attention is primarily attracted not to the *tactus* but to the melodic-harmonic material near the musical surface.

The segmentation of the middle level is primarily determined by a combination of (or interaction between) two factors: the duration of the segments and gaps between them, and the thematic structure. Both of these are considered in the segmentation rules of Lerdahl and Jackendoff (1983); non-durational aspects such as repetition, cadence, and dissonance-resolution can influence phrasing when they are incongruent with durational patterns (Fraser, 1993). As an example of the role of non-durational elements, consider the opening of the well known D-major "Ode to Joy" theme in Beethoven's 9th Symphony. A plausible account of Schoenberg's generative motive in this case might be "repeated note plus rising scalar third in four evenly-spaced onsets". The first 4-bar phrase would then include the motive plus three variants. But why do we parse this as a 4-bar phrase and not as two 2-bar phrases? These 4 bars form the first unit of a statement-plus-response archetype, itself forming the first part of a four-phrase archetype in which bars 9-

12 compact repetition-durations and prepare a return (statement, response, divergence, return). If the middle level is complex (see below), it may consist of *both* the 2-bar *and* the 4-bar phrase, and their interrelationship.

4. TOPICS

The most salient aspect of music for the listener may not be the motivic patterns in pitch-time space *per se*, but those aspects of the human cultural environment that the musical patterns refer to: "physical source, physical space and proximity, genre, musical function, performance skill, emotional attributes and social context" (Dibben, 2001, p. 161). For example, British listeners hearing a performance of Elgar's "Land of Hope and Glory" will immediately be aware of its cultural associations – British nationalism, the Empire, Queen Victoria – and associated emotional responses. The cultural function of a musical theme corresponds to its *affordance* in ecological psychology (Gibson, 1979). The affordances (e.g., uses) of a physical object are generally more salient to an observer than its physical properties. Similarly, the affordances of a musical fragment (meanings) are more salient to the listener than structural features such as contour and rhythm (cf. Clarke & Dibben, 1997); consider for example the archetypal 18th-century phrase whose frequency of occurrence in the classical musical literature was investigated by Gjerdingen (1984).

This leads us to consider the idea of *topics*, introduced by Ratner (1980) in his historical and style-critical treatment of classical music, and absorbed into the semiotic music-analytic paradigm of Agawu (1991). Topics are passages of music within larger pieces that refer directly or indirectly to specific musical-cultural entities (styles, moods, meanings, conventions) such as "pastoral", "romantic symphony" or "Turkish music as understood in 18th-century Vienna". Topics change rapidly in some styles (e.g., Mahler) and slowly in others (sBach). They are laden with expressive meanings that contribute to the musical discourse.

Musical topics exist, and seem to be perceived, primarily at the middle level; in the context of a musical piece, it typically takes a few seconds to realise that a new topic has arrived, to recognize it, and to become aware of its various associations. Familiarity with the motivic structures that typically generate a given topic may influence how the middle level is parsed (i.e., segmented and functionally interpreted by pattern matching). The analogy between topics in everyday speech and topics in music invites a new kind of psychological investigation of the relationship between music and language: topics may in both cases be matched to the knowledge and expectations of the listener, to maximize communicational efficiency. In linguistics, this involves using an appropriate vocabulary or *register* (cf. Eggins & Martin, 1997).

The idea of topics covers only a part of the more elaborate area of style- and generic-concepts. Genre and style promulgate expectational horizons or categories enabling meaning generation, and stylistic quotation and modulation form significant parts of the narrative trajectory of many pieces. We are interested in focal attention and the deployment of appropriate schemata in response to stimulus (such as military

fanfares in Mahler symphonies), including changes of schema promoted by strong triggers, such as the start of the pastoral second subject of the first movement of Brahms's Third Symphony (Pascall, 2001). We are also interested in the effect of topic on the way music is segmented into meaningful chunks, a process characteristic of the middle level.

Because different hierarchical levels interact and intertwine, the middle level, like any other level of analysis, may itself be *complex* and include secondary segmentations. It may also be multiple in the sense that different listeners may perceive more than one (adjacent) level simultaneously, or ambiguous in that their attention jumps back and forth between levels. From an empirical psychological point of view, there may be no clearly operationalized difference between multiplicity and ambiguity. Intuitively, we suppose that hierarchically adjacent candidates for the middle level can be folded into each other, giving the middle level complexity, substance, and teleological direction.

5. PSYCHOLOGICAL EVIDENCE

The idea of a middle level of musical experience and analysis is related to Fodor's (1983) exegesis of category hierarchies. Fodor showed that, in everyday language, salience "clusters at the middle levels of abstraction". We more often talk about "dogs" than about "carnivores", "mammals" and "vertebrates" on the one hand, and "hunting dogs", "hounds", and "blood hounds" on the other. Out of the *implication hierarchy* of possible formulations, we choose one somewhere in the middle, where the level of detail is neither too great nor too small for our communicative purpose. Fodor unpacked a default category of perception operating under normal conditions; the middle level is chosen for its inherent and transparent meaningfulness, for its efficiency in negotiating the world and in enabling communication, and for its organizing and cohering properties within its implication-hierarchy. In a given context, the listener takes advantage of the available information to determine the intended meaning and resolve potential ambiguities (Miller, 1999).

Fodor points out that there is no independent definition of "middle" – "and it is quite conceivable that intuitions about which levels are in the middle just are intuitions of relative salience". He then attempts to break this circularity by an analysis of characteristics of a middle level ("dog"): it has a high frequency in vocabulary counts, and is learnt earlier than other levels. Both these points have a musical analogy. First, as Schoenberg and Ruwet showed, musical phrases recur more or less frequently in different parts of a work, in different works and even in different styles. Second, children learn to understand and create both speech and music by actively perceiving, mimicking and improvising fragments of infant-directed speech (Papousek, 1996). The phrase level is privileged very early – infants in the study of Jusczyk and Krumhansl (1993) were remarkably sensitive to the phrase structure of Mozart minuets. It then continues into childhood: school children are good at judging the similarity of successive musical phrases (DeNardo & Kantorski, 1995).

The average duration of a musical phrase corresponds to that of a phrase in speech, which in turn corresponds roughly to the

period of relaxed breathing – typically about 4 seconds (Hildebrandt et al., 1998). Intuitively, this corresponds to the most salient level at which meaning is extracted from language: the intended meaning of a word is not clear until it appears in the context of a clause (with subject and verb). The duration of a phrase also corresponds roughly to the *psychological present* as defined by Fraisse (1963) and Michon (1978). The psychological present may be regarded as a level of memory that enables the different parts of a short sentence to be related to each other, enabling the utterance to be understood. In music, the psychological present enables the events in a phrase to be related to each other. Its duration depends on task (Crowder, 1993), just as the duration of perceived segments at the middle level depends on tempo.

The psychological present is related to Baddeley's (1986) concept of *working memory* with its emphasis on active, dynamic cognitive processing. While listening to a piece of music – or attending to any temporal stimulus, for that matter – we process it in working memory. Initially, we can only process those fragments of the music that can easily be stored in the working memory buffer. Cognitive processing of this kind may be regarded as a prerequisite for entry to longer-term storage, which in music enables topics to be recognized.

6. MUSIC-PSYCHOLOGICAL EVIDENCE

A range of music-psychological strands of research are consistent with the concept of the middle level. Perhaps the clearest piece of evidence is that familiar melodies can generally be recognized within the space of a few tones (Croonen, 1994) on the basis of their contours and rhythms (Dowling, 1978).

During the course of a melody, expectations change with every new note, as does the strength and clarity of expectations, which are relatively clearly defined at the middle of a phrase (Boltz, 1993). Psychological theories of melodic expectancy are largely confined to melodic fragments (Krumhansl, 1995) and so correspond to the middle level. Applying Meyer's (1956) theory of musical emotion, we might therefore expect emotional implications to be relatively intense at the middle level, consistent with Sloboda's (1991) finding that melodic appoggiaturas are important music-structural triggers of emotional response.

Tillmann and Bigand demonstrated that local rather than global musical contexts – corresponding to the middle level, as defined here – are sufficient for the recognition of a musical fragment (Tillmann & Bigand, 1998), predominate in the cognitive processing of harmonic cadences (Tillmann et al., 1998), and contain enough information to define the music's expressiveness (Tillmann & Bigand, 1996). On this basis, the musical surface deserves to receive more attention, in both music theory and music psychology (Kreutz, 1998). The idea of middle-out analysis is a step in that direction.

7. REFERENCES

- Agawu, V. K. (1991) *Playing with signs: A semiotic interpretation of classic music*. Princeton Univ. Press.
Baddeley, A.D. (1986). *Working memory*. Oxford U. Press.

- Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics*, 53, 585-600.
- Cambouropoulos, E., & Widmer, G. (2000). Automated motivic analysis via melodic clustering. *Journal of New Music Research*, 4, 303-317.
- Clarke, E., & Dibben, N. (1997). An ecological approach to similarity and categorisation in music. In M. Ramsar et al. (Eds.), *Proceedings of the Interdisciplinary Workshop on Similarity and Categorisation* (pp. 37-41). Department of Artificial Intelligence, Univ of Edinburgh.
- Croonen, W. L. M. (1994). Effects of length, tonal structure, and contour in the recognition of tone series. *Perception & Psychophysics*, 55, 623-632.
- Crowder, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound* (pp. 113-145). Oxford: Clarendon.
- DeNardo, G. F., & Kantorski, V. J. A. (1995). A continuous response assessment of children's music cognition. *Bull. of the Council for Res. in Music Education*, 126, 42-52.
- Dibben, N. (2001). What do we hear, when we hear music?: Music perception and musical material. *Musicae Scientiae*, 5, 161-194.
- Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85, 341-354.
- Drabkin, W. (1985). A lesson in analysis from Heinrich Schenker: The C major Prelude from Bach's Well-Tempered Clavier, Book I. *Music Analysis*, 4, 241-58
- Eggins, S., & Martin, J. R. (1997). Genres and registers of discourse. In T. A. van Dijk (Ed.), *Discourse as structure and process: A multidisciplinary introduction*, Vol. 1. (pp. 230-256). Thousand Oaks, CA: Sage:
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fraisse, P. (1963). *The psychology of time*. New York: Harper & Row.
- Fraser, R. (1993). Computer perception of phrase structure. *Contemporary Music Review*, 9, 255-266.
- Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gjerdingen, R. O. (1984). *A musical schema: Structure and style change, 1720-1900*. PhD diss., U. of Pennsylvania.
- Hildebrandt, G., Moser, M., & Lehofer, M. (1998). *Chronobiologie und Chronomedizin*. Stuttgart: Hypokrates.
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *J. of Experimental Psychology: Human Perception and Performance*, 19(3), 627-640.
- Koch, H. C. (1782-93). *Versuch einer Anleitung zur Komposition*. Leipzig: A. F. Bohme.
- Kreutz, G. (1998). *Musikalische Phrasierung aus historischer und kognitionspsychologischer Sicht*. Frankfurt: Lang.
- Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*, 17 (1), 53-80.
- Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT.
- Meyer, L.B. (1956). *Emotion and meaning in music*. University of Chicago Press.
- Meyer, L.B. (1967). *Music, the arts, and ideas*. University of Chicago Press.
- Meyer, L.B. (1973). *Explaining music: Essays and explorations*. University of Chicago Press.
- Michon, J. A. (1978). The making of the present: A tutorial review. In J. Requin (Ed.), *Attention and performance* (Vol. 7). New York: Academic.
- Miller, G. A. (1999). On knowing a word. *Annual Review of Psychology*, 50, 1-19
- Narmour, E. (1977). *Beyond Schenkerism*. University of Chicago Press.
- Nattiez, J. J., transl. A. Barry (1982). Varèse's "Density 21.5": A study in semiological analysis. *Music Analysis*, 1/3, 243-340.
- Papousek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliège & J. Sloboda (Eds.), *Musical beginnings* (pp. 88-112). Oxford U Press.
- Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11, 409-464.
- Ratner, L. (1980). *Classic music: Expression, form and style*. New York: Schirmer.
- Pascall, R. (2001). Le processus compositionnel à la lumière d'un procédé analytique de type sémiotique dans la 3ème symphonie de Brahms. In *Analyse et création musicales: Actes du Troisième Congrès Européen d'Analyse Musicale* (pp. 459-483). Paris: Éditions l'Harmattan.
- Rothstein, W. (1989). *Phrase rhythm in tonal music*. New York: Schirmer.
- Ruwet, N. (1966). Méthodes d'analyse en musicologie. *Revue Belge de Musicologie*, 20, 65-90.
- Schenker, H. (1935): *Der freie Satz*, transl. E. Oster (1979): *Free composition*. New York: Longman.
- Schoenberg, A. (1943). *Models for beginners in composition*. New York: Schirmer.
- Sloboda, J. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, 19, 110-120.
- Tillmann, B., & Bigand, E. (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, 24(1), 3-17.
- Tillmann, B., & Bigand, E. (1998). Influence of global structure on musical target detection and recognition. *International Journal of Psychology*, 33(2), 107-122.
- Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*, 61(3), 157-174.