The Butterfly Schema as a Product of the Tendency for Congruence and Hierarchical Selection in the Instrumental Musical Grammar of the Classical Period

Trevor Mark Rawbone

A thesis submitted to the University of Huddersfield
in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

March 2017

(Word Count: 80,267)

Abstract

Diverging explanations of local multiparametric schemata are found in music of the common practice period (c. 1600–c. 1900). Associative-statistical theories describe schemata as situated structures in particular times and places, whereas generative theories present these constructions as features formed through stability in universal and general rule systems. Associative-statistical theories of schemata elucidate the culturally conditioned relationships between features (distinctive attributes commonly used in grammars and schemata), but do not show the influence of universal psychological constraints; generative theories reveal the implicit structure of music, but do not formalise particular grammatical features and contexts. A synthesis of generative and associative-statistical approaches is necessary to model the interaction between universal and particular constraints of grammars and schemata. This dissertation focuses on a novel localised schema formed in the Classical instrumental grammar, termed the butterfly schema. It is posited that the butterfly schema is generated by a tendency for congruence that is manifest in and between the particular features of this grammar.

Computational musicology and psychology provide interdisciplinary insight on the formal possibilities and limitations of grammatical structure. Computational models of schemata and grammars show how the congruent features of musical structure can be represented and formalised. However, they also highlight the difficulties found in the automatic analyses of multiparametric relationships, and may be limited on account of their inductive frameworks. Psychological approaches are important for establishing universal laws of cognition, but are limited in their potential to account for the diversity of musical structuring in grammars. The synthesis of associative-statistical and generative approaches in the present dissertation permits modelling the combination of the universal and particular attributes of butterfly schemata. Butterfly schemata are dependent on the particular grammars of periods of history, but are constrained by the tendency for congruence, which is proposed to be a cognitive universal. The features of the butterfly schema and the Classical instrumental grammar are examined and compared against the features of the Baroque and Romantic grammars, showing how they are formed from diverse types of congruent structuring. The butterfly schema is a congruent grammatical category of the Classical instrumental grammar that comprises: chords that are close to the tonic in pitch space (with a chiastic tension curve starting and ending on the tonic); a textural and metrical structure that is regular and forms a regular duple hierarchy at the level of regular functional harmonic change and at two immediately higher levels; and simple harmonic-rhythm ratios (1:1 and 3:1).

A survey conducted using arbitrary corpora in European instrumental music, *c.* 1750–*c.*1850, shows the distribution of butterfly schemata. Butterfly schemata are more common in the Classical-period sample (*c.* 1750–*c.* 1800) than in the Romantic-period sample (*c.* 1800–*c.*1850), suggesting that the tendency for congruence manifest in and between the features common in the Classical grammar generates butterfly schemata. A second component to the statistical analysis concerns the type of schemata observed, since the tendency for congruence is presumed to also apply to the type of features that form in butterfly schemata. Maximally congruent features are generated more commonly than minimally congruent features, indicating the influence of the tendency for congruence. This dissertation presents a formulation of the Classical instrumental grammar as a multiparametrically congruent system, and a novel explanation and integration of the concepts of grammars and schemata. A final component to the dissertation poses that the features of the Classical instrumental grammar and butterfly schema follow a distinct order of dependency, governed by the mechanism of selection in culture. Although the tendency for congruence governs all features of a grammar, features are also formed by the top-down action of culture which selects those features. Thus, a top-down hierarchical selection model is presented which describes how the butterfly schema is formed through the order of selection of features in the Classical instrumental grammar.

Table of Contents

List of Figures

Dedications and Acknowledgements

I would like to thank my supervisor, Dr. Steven Jan, for his untiring patience, kindness, and willingness to permit me freedom to explore various avenues of music theory, while also providing the necessary sanity checks to keep me on course. I am grateful for his help in procuring teaching opportunities and for jointly authoring our articles. I am also grateful to my second supervisor, Prof. Michael Russ, for his wise input into this project during the early stages. I thank the School of Music, Humanities and Media at Huddersfield University for enabling this project to take place and providing a scholarship towards this research. Also, Cardiff University Music Library has been very helpful in providing resources during my years of study in Wales.

I am indebted to my grandparents who have supported me over the course of this research, and would like to dedicate this dissertation to the memory of my grandfather, who left this world just a few weeks before its completion. My brother has been a constructive critic of my work, and, with utmost tact and sensitivity, incentivised many refinements of ideas. My piano teachers, Mr. Mark Shewring and Prof. Michael Young, deserve credit for their wisdom and education, which I hope even a little might have been instilled in me and found its way into this dissertation. I thank my wife Khallud Saigh for her support and understanding which have encouraged me and allowed this work to be carried out much more smoothly.

Chapter 1: Introduction

This dissertation presents the butterfly schema as a novel model of a distinct localised parcel of multiparametric features. The butterfly schema comprises a chiastic chord structure, starting and ending on the tonic, with congruent features of texture, harmonic rhythm, and metrical and tonal structure. The appellation 'butterfly' characterises the symmetrical chord structure and hints at the expressive lightness felt through its stable structure. It is found in a number of grammars of various musical periods, including 'light' music and opera of the nineteenth-century, in the music of composers such as Vincenzo Bellini (1801–1835) and Guiseppe Verdi (1813–1901) and popular music of the early-twentieth century, but is here shown to be common in compositions of European instrumental music of the Classical period. In broad terms, grammars are global rule systems of musical structure, and schemata are local statistical regularities of musical structure. In the various sub-disciplines of music theory, grammars and local schemata receive varying explanations (e.g., Meyer (1973), Lerdahl and Jackendoff (1983), Gjerdingen (1988), Leman (1995), Lerdahl (2001), Temperley (2004), Kaiser (2007b), Byros (2009a)). Understanding the relationship between the Classical instrumental grammar and the butterfly schema is a primary aim of this dissertation.

The present examination of the Classical instrumental grammar and butterfly schema draws together theories of musical structure from diverse subfields, including computational grammars (Longuett-Higgins and Steedman, 1971; Tenney and Polansky, 1980; Hamanaka et al., 2005), Gestalt psychology (Wertheimer, 1923), auditory stream analysis (Bregman, 1990), behavioural psychology (Skinner, 1953, 1981), memetics (Dawkins, 1976; Jan 2007, 2013), as well as the core discipline, music theory (Cooper and Meyer, 1960; Cone, 1968; Gjerdingen, 1988; Lerdahl and Jackendoff, 1983; Lester, 1986; Rothstein, 1989, 1995; McKee, 2004). Theories of schemata can be broadly classified into one of two categories, being formed either as a product of generative rules or through the statistical association of schema features. Generative theories show how schemata form through stability in grammars, through a combination of cognitive universals and cultural constraints (as claimed in Lerdahl and Jackendoff (1983, p. 289) and Temperley (2001, pp. 336–340)), whereas associative-statistical theories explain schemata primarily as statistical regularities in culture that are abstracted by cognition (as argued by Gjerdingen (1988, p. 99) and Byros (2009a)). These require reconciliation for understanding the causes of the butterfly schema.

A synthesis of generative and associative-statistical theories of schemata becomes possible through the notion of the *tendency for congruence*, which is the inclination for grammars and schemata towards stability in and between their parametric features. Congruence in musical structure may be described as stability or agreement in and between the features of grammars and schemata, such as when the elements of a chord are consonant, or when stable chords occur with stable metrical accentuation. The tendency for congruence is proposed to be a preference of cognition that manifests itself in and between the cultural features of grammars, forming local schemata. A comparison of the Classical grammar[1] with other grammars, such as those of the Baroque and Romantic periods, shows how the tendency for congruence is particularly manifest, and explains how the butterfly schema is generated. The tendency for congruence is an indirect product of cognition and its effects can be measured empirically. This concept is supported through analysis and a statistical survey to show the commonality of the butterfly schema as a product of the Classical grammar.

Butterfly schemata are also argued to emerge because culture selects particular features in certain times and places, while being constrained by the tendency for congruence. A hierarchical model of cultural selection is therefore introduced to show the interaction of features in the Classical grammar and butterfly schema. These features are cumulatively selected in a top-down manner. Through a model of cultural selection, the butterfly schema is elucidated as a structure formed through the universal tendency for congruence in cognition and the particular culturally selected features of the Classical grammar.

---

[1] For brevity, the 'European Classical instrumental grammar' is generally shortened to 'Classical grammar' throughout this dissertation.

This chapter provides an overview of how this dissertation supports the notion that the tendency for congruence (in and between the features of the Classical grammar) and top-down hierarchical selection produces butterfly schemata. The dichotomy between associative-statistical and generative theories of schemata is reviewed. Both are shown to be insufficient approaches when used in isolation for explaining the causes of butterfly schemata. The tendency for congruence provides a way to reconcile these approaches, constraining the generation of the butterfly schema in the Classical grammar. It is a universal product of cognition that governs musical structure. However, it cannot be used to explain cognition, but is an explanation of the effect of cognition on musical structure. Therefore it is necessary to show that while this conception assumes rationalist metaphysics, it is not provide a theory of inductive inference. A rationalist explanation of musical cognition contends that the implicit congruent and noncongruent interrelationships in grammars and schemata are *a priori* concepts, intuitively understood by listeners, not measured through enumeration. To this end, this chapter includes a review of the philosophical, psychological, and cultural foundations of grammars and schemata in order to provide the reader with an entry point into the issues of congruence. (Section 1.2 explores the philosophical, psychological, and cultural influences on schemata, and Section 1.3 examines how these frame the tendency for congruence.)

1.1 Associative-Statistical and Generative Theories of Schemata

Associative-statistical and generative theories of schemata provide a foundation for dealing with the concept of the tendency for congruence and the butterfly schema. This section provides an introductory examination of these theories, providing a brief explanation and suggesting a possible synthesis. Chapter 2 goes on to examine associative-statistical and generative theories in greater depth, while Chapters 4 and 5 develop a comprehensive theory of the tendency for congruence. A staggered explanation of this concept over these chapters is necessary because it requires detailed explanation in uniparametric and multiparametric contexts. Also, a background of associative-statistical and generative theories (covered in

Chapter 2) and an understanding of the psychological, computational, and cultural theories (provided in Chapter 3) are necessary before the butterfly schema can be fully formalised.

1.1.1 Associative-Statistical Theories of Schemata

Associative-statistical theories pose that schemata are cognitively internalised through the statistical association of networks of features in particular times and places (e.g., Meyer (1973), Gjerdingen (1988, 2007), Byros (2009a)). The presentations of schemata in Gjerdingen (1988, 1996, 2007) rest on the statistical relationships between features of schemata in a particular time and culture. Schemata exist in cognition on account of regularities in the environment (Gjerdingen, 2007, p. 11). However, principles of perception and cognition seem to constrain how these are internalised. Gjerdingen's (1988, 2007) definition of schemata rests on the perception of distinct voice-leading patterns (or 'events') in the melody. A number of perceptual and cognitive principles are required to extract and coordinate features. Voice-leading events are perceived and cognised through principles such as clarity, distinctness, vividness, and prominence (Gjerdingen, 1988, 85). The definition, validation, and recognition of schemata is made with reference to an ideal prototype, which is abstracted from a continuum of historical change (Gjerdingen, 1988, 99). The popularity and typicality of the 1–7…4–3 schema appears to rise and fall in history, exhibiting a Guassian distribution that peaks in the early 1770s. Human cognition (which is presumably universal) abstracts similar types of structures in the historical continuum (Gjerdingen, 1988, 99).

The abstraction of schemata from the environment implies that perception and cognition is important for defining schemata. However, in Gjerdingen (1988) and Byros (2009a) there is arguably no systematic framework to explain how this is done (excepting the broad notions of clarity, distinction, vividness, and prominence). Moreover, there is not an adequate description of the conditions for defining the prototype, the process of abstraction, the methods of validation, and the causal mechanism of how schemata form. Also, there are arguably no systematic frameworks or methods for explaining why features should be coordinated in a prototypical way (Cavett-Dunsby, 1990; Cohn and Dempster, 1992). Furthermore, an explanation of whether schemata might be a product of broader principles of cognition and grammar is not provided. Thus, associative-statistical theories are perhaps more empiricist explanations, rather than cognitive or rationalist theories, since they establish that culture is fundamental for the formation of schemata.

1.1.2 Generative Theories of Schemata

Generative theories of music depict global rule systems of a corpus that are shaped by universal psychological constraints. Generative grammars can also be defined as formal descriptions of the musical intuitions of listeners experienced in a particular idiom (Lerdahl and Jackendoff, 1983, p. 1). Generative theories often view schemata as the conglomeration of stable features in a local context, such as tonic or dominant harmony, duple phrasing, and parallel metrical structure, etc. In *A Generative Theory of Tonal Music* (*GTTM*) (Lerdahl and Jackendoff, 1983), schemata emerge due to stability in the system of well-formedness and preference rules (Lerdahl and Jackendoff 1983, p. 289). In *The Cognition of Basic Musical Structures* (*CBMS*) schemata are 'occasional' structures that are dependent on the 'ubiquitous' infrastructure of grammar (Temperley, 2001, pp. 3–4).

Tonal Pitch Space (*TPS*) (Lerdahl, 2001, pp. 233–242) explains schemata as a product of the prolongational, harmonic, grouping, and metrical tendencies in structure that are constrained by cognition. However, Lerdahl (2001, p. 248) points out that schemata also represent an emergent level of analysis, suggesting that they have elements that are more specific than can be explained through grammatical determinism, requiring further cultural constraints (as argued in Gjerdingen (1988) and Byros (2009a)). Thus while generative theories of schemata combine the universal generative rules of cognition with the rules of the tonal idiom, they generally do not provide insight into the fine-grained sub-cultural styles of tonal music, as do associative-statistical approaches (e.g., Gjerdingen (1988)), and therefore might present too generalised readings of tonal music that can be incoherent for particular grammars.

1.1.3 A Synthesis of Generative and Associative-Statistical Theories of Schemata

In broad terms, schemata are particular structures in associative-statistical theories (Meyer, 1973; Gjerdingen, 1988; Byros, 2009a), but generic and universal structures in generative theories (Lerdahl and Jackendoff, 1983, p. 289; Lerdahl, 2001, pp. 231–242; Temperley, 2001, pp. 336–340). Statistical-associative models describe historically and geographically situated schemata, but do not explain the universal grammatical constraints that govern their formation. By contrast, generative models are concerned with the universal generative capacity of music, in the context of generic tonal idioms, but do not explain the emergence of particular subcultural features that emerge in specific times and places.

For Rohrmeier and Neuwirth (2015, p. 290), the dichotomy between associative-statistical and generative theories of schemata hinges on the generality of the structure in question: '[i]f a pattern evinces both regularity and combinatorial freedom, grammars will be more suitable to describe it; however, if a musical structure exhibits more standardisation and less variability, schema-theoretical (and exemplar-based) approaches will be more appropriate'. Notwithstanding, the ontological and epistemological status of localised schemata (such as the butterfly schema in the Classical grammar) require deeper analysis, having aspects that are both stable and combinatorial, being a product of both universal generative psychology and particular cultural conditioning. Therefore an explanation of schemata requires a synthesis of associative-statistical and generative explanations. A synthesis is possible through the tendency for congruence.

The tendency for congruence is a universal propensity of cognition, acting on the particular features of the Classical grammar, generating congruent parcels of particular local features, such as butterfly schemata. It constrains the uniparametric internal structure of features, and also their multiparametric interaction, causing a high frequency of multiparametrically congruent local schemata. It permits a more coherent explanation of the interrelationships between universal and particular constraints, abstract and concrete levels, and local and global structure than generative or associative theories used in isolation. However, it is first necessary to explain the philosophical principles underlying the tendency for congruence (examined in Section 1.2) before it can be more fully outlined in Section 1.3.

1.2 Philosophical, Psychological, and Cultural Foundations of Grammars and Schemata

In this section, the non-inductive, rational, and selective notions that underpin this theory of grammars and schemata are presented. The tendency for congruence presupposes, but does not explain, a rationalist explanation of human understanding of musical structure. It does not provide a cognitive explanation of musical structure that is based on inductive inference because empiricist methodologies cannot explain cognition. This is argued with reference to the philosophical notion of the problem of induction. The tendency for congruence merely explains a statistical trend as a product of cognition. However, empiricist, associationist, selectionist, and inductive theories are also required in this theory of grammars and schemata for a complete understanding of musical structure since they account for the influence of culture.

Broadly, generative theories of schemata are rationalist and deductive, concerning internal cognitive principles, and are also qualitative, being based on broad mentalistic interpretations of structure. Furthermore, generative theories are abstract and universal since they involve systematising the generalised structures of cognition and culture. By contrast, associative-statistical models are empiricist and inductive, concerning the interaction between culture and behaviour. They are quantitative and statistical, being data-driven interpretations. Also, associative-statistical models are concrete and particular, involving the specific traits in cultures. The model of the butterfly schema is proposed to integrate these well-entrenched divisions between generative and associative-statistical explanations. The tendency for congruence is a constraint of cognition that interacts with the associative and selective mechanisms of culture. This section explores this interaction between universal cognitive constraints and particular cultural constraints, focussing on the abstract features of tonal and metrical structure.

1.2.1 The Problem of Induction

The rationalist perspective on human understanding is epitomised in Descartes' (1637/1978, p. 27) epithet, *I think, therefore I am*. Since the rational observer is able to contemplate, and indeed question his own existence, he must be separate from the material being contemplated. The conscious observer (located in the pineal gland according to Descartes) is therefore distinct from the material world. By contrast, empiricist philosophy views this dualistic position as an idealist fiction (a construct of minds) that circumvents the problem of explaining the causal mechanism of cognition, and which enables the appearance of intelligence and understanding (Pagel, 2012, p. 270).[2] Locke (1689/1993, pp. 231–240), leaning towards nominalist metaphysics (where general ideas exist only in minds but not in actuality), criticises human knowledge as digressions of verbal epistemology (as viewed by Russell, 1946, p. 556). According to Locke (1689/1993, pp. 45–224), knowledge is a complex web of associations, assumptions, and approximations constructed through sense data, and so human understanding has few innate ideas, principles, or points of coherence.

Locke's (1689/1993) empirical explanation of human understanding, via sense data, is useful for explaining the conditioning effect of culture, but does not explain how humans use cognition to organise the world, and further, how they coherently act on knowledge. Hume (1739–40/1985, pp. 135–142) nuances the debate between rationalism and empiricism through what is now most commonly referred to as the problem of induction. Hume notes that on first approximation, knowledge of causal relations in the world seem to be inductively inferred though probabilistic relations between objects. Hume argues that the knowledge that fire causes pain is not rationally understood through deductive reasoning in cognition, but is induced through the experience of the effect of fire on the human body. Inductive inference thus appears to establish knowledge about causal relations between objects through statistical enumeration, not critical deduction. However, Hume questions whether knowledge can solely be based on experience this way, precisely because it cannot provide a causal explanation of

---

[2] Pagel (2012, p. 270), drawing from Hume (1739), provides a succinct refutation of the so-called mind-body problem (although it is not completely satisfactory because rationalism seems to be intuitively true): '[W]e now know that Hume's instincts were right: there is no central place where our brains collect up all our thoughts and present them for "you" to watch. If there were an inner "you", then that would also have to have an inner you, and so on, ad infinitum'.

the interaction between objects (Hume, 1739/40/1985, pp. 135–142), which must be necessary for many forms of human understanding. Likewise, Popper (1935, pp. 3–20, pp. 253–254) argues that the establishment of knowledge or a theory based on empirical evidence is a flawed notion since evidence is finite and therefore does not provide justification of those notions. Thus theories based on inductive inferences are almost always false. For example, the number of white swans observed does not justify a conclusion that all swans are white, since the existence of a black swan would show that this inference had no logical basis (Popper, 1935, pp. 3–4). Knowledge cannot be gained through inductive inference, or any other quantitative method, but must be based on rational deductive criticism (Popper, 1935).

It has been argued that knowledge must be abductively understood in cognition, since inductive inference cannot form a basis for understanding. A corollary of this argument is that statistical tendencies that can be observed in musical structure do not provide an explanatory framework for cognition. This non-inductive conception assumed in the tendency for congruence is examined in more detail in Section 1.3 (examined in Section 5.3).

1.2.2 Cognitive versus Cultural Explanations of Musical Structure

This section shows that although universal cognitive principles form the core axioms of generative theories, they require reconciliation with particular cultural constraints. The generative enterprise in linguistics, propounded by Chomsky (1957, 1965), is a foundational framework of many generative approaches in music theory. Chomsky (1957) posits that humans must have a built-in capacity to generate well-formed grammatical structure. Well-formed grammars are constructed from the constituent syntactic rules used to generate them. Cognition regenerates (*wiedererzeugt*) (Chomsky (1975, p. 93), after Humboldt (1836/1999, p. 93))[3] language structure, and presumably musical structure also. Cognition regenerates well-formed structure with relatively little prior exposure to language and music, described as the 'poverty of the stimulus' argument (expounded in Chomksy (1965, pp. 27–30)). However,

[3] Schenker (1935/1979, p. 6) also argues for a universal cognitive framework for understanding music: 'Nothing new is to be expected, yet this need not surprise us when we see that even in technology, which today stands in the forefront of all thought and activity, nothing truly new appears: we witness only further transformations.' However, many well-formed musical styles are not revealed through Schenkerian theory because it is idiom-specific (Narmour, 1977).

grammars are also governed by cultural constraints, and so to provide a comprehensive model of grammars and schemata it is necessary to account for these. In the present dissertation this is achieved through considering the cultural selection of features in a grammatical hierarchy. A top-down hierarchical selection (HS) model shows the cumulative selection of features by the top-down action of culture (see Section 5.3).

A difficulty for modelling grammars and schemata is how to distinguish between the universal constraints and the particular features of musical structure. At first blush, the common structures between musical cultures might seem to be universal and the non-common structures particular. However, this distinction is misleading because similar structures can have diverse roles in different contexts (Meyer, 1956, pp. 45–50). Establishing the universal causes of musical structure requires intuitive understanding of the implicit congruent structure. There is no systematic method to show how this might be done, although it seems to require the rational powers of cognition. Moreover, distinguishing between the universal and the particular is perplexing because regenerated structure can be considered both universal and particular. That is, particular features correspond with universal principles because psychologically preferred structures (which are regenerated by cognition) emerge in various grammars. Conversely, while the universal constraints of grammars govern particular features, they must also be selected by culture in certain times and places. These conceptions blur the distinction between the universal and particular in musical structure, and show that understanding differences requires a fine-grained analysis of specific contexts.

Aspects of the Classical grammar, such as its common chord progressions, can be described as both universal and particular. The harmonic series acts as a broad universal psychoacoustic constraint on the cognition and form of pitch and harmony (Rameau, 1722; Helmholtz, 1863/1875; Huron, 2001; Lerdahl, 2001). Lerdahl (2001) presents a number of basic spaces (diatonic space, octatonic space, hexatonic space, etc.), which are universally constrained ways of hearing pitch relationships in terms of chromatic, diatonic, triadic, fifth, or octave levels. These basic spaces are governed by a number of universal 'constraints on basic spaces' (Lerdahl (2001, pp. 268–274) (Section 2.3.7 more fully examines these constraints). The distances between chords, pitches, and keys are universally conceived in terms of the diatonic, triadic, fifth, or octave relationships of these basic spaces (Lerdahl, 2001). Therefore seemingly culturally particular chord progressions of the Classical grammar are actually regenerated by cognitive universals, since they are governed by relationships in diatonic basic space, which are, in turn, regenerated from the universal constraints on basic spaces. Conversely, both chord progressions and the basic spaces could be viewed as particular constructions because they are unevenly distributed in the world, transmitted between people and artefacts in culture.[4] That is, universally constrained features are selected in certain musical traditions, histories, and geographies, and not in others, shown in the top-down HS model (presented in Section 5.3).

The Gestalt 'laws' of proximity, similarity, and good continuation (Wertheimer, 1923; Koffka, 1935) form some of the primary universal cognitive constraints that govern the way both rhythm and pitch structure is perceived. However, while Gestalt psychology sheds light on the principles of perception, it is also the case that musical structure is not fully governed by Gestalt principles. Incorporating Gestalt principles of proximity and similarity as foundational laws, as basically proposed in many generative theories (such as *GTTM* and *CBMS*), cannot account for the breadth of variance in musical structure, even within tonal music. It is therefore preferable to term universal influences on grammars 'constraints',[5] rather than 'rules' (as suggested by Lerdahl, 2012), since differences between grammars rest not on the presence or absence of universals, but how these interact with particular constraints. For example, a universal constraint on phrases is that they are close to an ideal duration and size due to the limits of short-term memory: between three and five seconds (Synder, 2000, p. 13) and between five and nine elements (Snyder, 2000, p. 36) (roughly eight notes in Temperley (2001, p. 69)). However, the phrase lengths in grammars can be longer or shorter than these universal constraints, or can contain greater or fewer elements. Such outcomes are not that universal constraints are not in effect in those grammars but that they interact with the particular constraints therein. So, as described, in addition to the particular features of grammars being governed by the universal regeneration of structure in cognition, grammars are a product of the selection of particular features.

---

[4] For example, diatonic space, which strongly adheres to the constraints of basic spaces in Lerdahl (2001, pp. 268–274), is particularly common in Western culture.

[5] 'Constraint' is used in Lerdahl (2001) and is broadly analogous with the term 'preference rule' of *GTTM*.

Due to its complexity, variability, and creativity, it is questionable whether a definitive explanation of musical cognition might ever be presented in music research due to limitations on the intelligibility of cognition in general (Chomsky, 2006). Therefore the claims of generative theories about musical cognition require close critical analysis. Generative theories of music, such as *GTTM*, *TPS*, and *CBMS*, are not wholly generative because they do not generate structure from concrete elements to produce well-formed structures (Rohrmeier and Neuwirth, 2015, pp. 296–297). Generative grammars in linguistics involve examinations of well-formed structure within a single hierarchy of syntactic relations. *GTTM*, *TPS*, and *CBMS* involve the interaction of multiparametric hierarchies of features which are of graded grammatical well-formedness. However, each parametric system, such as harmony, rhythm, or metrical structure, can feasibly form a generative recursive system when considered in isolation (as Rohrmeier (2011) does for harmonic structure). This provides a significant hurdle for a generative theory of musical structure, because it must necessarily incorporate graded multiparametric relationships. Moreover, it is likely that the structure and interaction of the parametric sub-systems or features of music vary from grammar to grammar. This means that grammaticality in musical structure is not just graded, but differentiated, which is not easily presented in systems of well-formedness. For this reason, the tendency for congruence is an important conception because it permits variable structural descriptions of music.

Associative-statistical theories are problematic not just because they do not present reliable accounts of musical cognition, but because they often provide limited explanations of the cultural features of particular rule systems. However, an associative or selectionist framework is necessary (depicted through a top-down HS model) to show how features of grammars and schemata condition other features, resulting in hierarchies of cumulatively selected features.

1.3 The Tendency for Congruence

This section provides an overview of a central theory of this dissertation, the tendency for congruence, proposed to act in and between the features of the Classical grammar to produce butterfly schemata. The discussion provides a foundation for an extensive examination of this notion in Chapters 4 and 5, after the incorporation of supporting theories of *GTTM* and *TPS*, among others. As discussed, the tendency for congruence describes the propensity for cognition to generate implicitly congruent structure, but does not provide an inductive theory of musical cognition. The tendency for congruence is a product of cognition that constrains the form and commonality of grammatical and schematic structure. However, in real-time listening congruence and noncongruence are *a priori* concepts that are intuitively understood, and so neither is preferred (the order of causation of features in real-time is further examined in Section 4.1.2). Notwithstanding, the ultimate causes of schemata in grammars are the tendency for congruence in cognition and the cultural selection of features (which act over a large time period in history).

1.3.1 The Role of the 'Rule of Congruence'

Longuet-Higgins and Steedman (1971, p. 223) employ a 'rule of congruence'[6] for automated tonic-finding and metrical analysis. It is used as a heuristic tool to parse the musical systems of time signature and key signature:

> If our rules are to be able to decide the time signature and key signature from the durations of the notes and their positions on the keyboard, some assumption must be made about how much of these data may safely be assumed congruent, and may therefore be used as evidence in reaching the required decision. The rule of congruence is such an assumption. It states that until a metric or harmonic signature has been established, every note will be congruent with it in the relevant attribute, unless the note is non-congruent with all other possible signatures. In other words, a non-congruence must not occur until it can be recognised as such. This is surely common sense. Music would be a dull affair if all notes had to be in the key and all accents on the beat, but it would be

[6] The 'rule of congruence' of Longuet-Higgins and Steedman (1971) is subtly different to Rothstein's (1995) use of this term to describe the non-inductive tendency for congruence between grouping and metre.

> incomprehensible if the key and metre were called into question before they were established' (Longuet-Higgins and Steedman, 1971, p. 224).

Longuet-Higgins and Steedman (1971) are arguing here that the rule of congruence explains how these musical structures are meaningfully inferred. In the tonic-finding model, templates of tonal collections, such as the key of C major, are compared against the pitch structure of passages to determine the key, in a pattern-matching process. According to Longuet-Higgins and Steedman (1971), dissonant notes outside the key are necessarily heard as noncongruent divergences from the template. Likewise, in the metrical analysis system, there are fixed templates of metrical structures that are used to infer metrical structure. Rhythms that contradict the metrical pattern are heard as diversions from metrical congruence. This conception presumes that music can only be understood with reference to a fixed pattern or it is incomprehensible. However, this cannot be an accurate explanation of how listeners understand music, since various degrees of congruence, including highly noncongruent relationships, are found in music. Such music must be still understandable, even though the patterns of tonal and metrical structure are idiosyncratic. Therefore such a rule of congruence cannot show how the knowledge of musical structure is inferred. As argued above, meaningful interpretations are not based on inductive inference. The implicit congruence and noncongruence of musical structure is intuitively understood in real-time listening by perception and cognition. How this is done is perhaps beyond the scope of present scientific understanding, but it can be assumed that it is done, and with relatively little prior exposure to the particular constructions of musical cultures.

That listeners in real-time have an immediate *a priori* understanding of structure can be seen, for example, through the direct knowledge that a major chord is intrinsically more congruent than a diminished chord, or that a isochronous metrical structure is more congruent than an non-isochronous metrical structure. This is achieved without reference to a theoretical framework, and without the application of previously learned schemata. Indeed, listeners understand musical structure even when it might contain a complex admixture of congruent and noncongruent elements. Lerdahl (2001, p. 153) describes how the chord structure of a passage, peppered with dissonant notes, must be conceived in terms of its abstract hierarchical harmonic structure. The accented passing notes E flat and C in the melody part of Figure 1.1 are conceived through abstracting the B-flat major chord shown in the lowest stave. Building on Lerdahl's analysis, it can be added that listeners must have an *a priori* understanding of the congruent interaction between the hierarchical systems of multiparametric features (the metrical structure, the grouping structure, the harmonic structure, etc.), because each must be interpreted in terms of the others. While listeners might have experience with these types of chord structures,[7] or might be familiar with similar rhythmic and metrical structures, the passage contains a novel arrangement of these congruent and noncongruent structures that is directly interpreted. Therefore this passage has an interaction between parametric features that is intuitively cognised, without the application of prior learning.

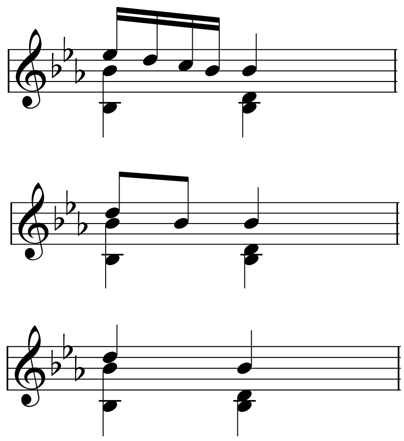


Figure 1.1 Hierarchical reduction of Mozart's Sonata in E-flat major, K. 282 (1774), i, bar 2 (Lerdahl, 2001, pp. 152–153).

While the chord and non-chord notes could be computationally identified using a template pattern for chords that are algorithmically checked against the passage, this type of heuristic tool cannot form part of a serious investigation into musical cognition because listeners must intuitively tease apart the interrelationships between the systems of pitch, rhythm, texture, and metrical and tonal structure. This rational process is a requirement because the degree of congruence in and between features is flexible and the patterns that form in music are infinite. The surface pattern was abstracted into the background pattern, not the converse. If the congruent interaction between interdependent domains was not intuitively recognised,

---

[7] Krumhansl (1990, pp. 240–270) observes that listeners internalise varied culturally conditioned schemata for tonal chords and keys, but there is general cross-cultural similarity in their representation and interpretation with repeated listening.

chord notes could not be distinguished from non-chord notes, and downbeats could not be separated from upbeats.

A final argument for the intuitive understanding of congruence, contrasting with the use of congruence as a heuristic tool of inductive inference (or part of a statistically learned schema), is that congruence is understandable even when it is not statistically frequent in a particular grammar. For example, although Wagner's late music generally avoids traditional progressions, often forming cadences that culminate on secondary triads, these are felt as diversions from an intuitive sense of closure that would have been experienced through concluding on tonic triads (Lerdahl, 2001, p. 176). That is, the veridical expectations of musical styles can conflict sharply with the universal and intuitive expectations generated through a direct perception and cognition of the implicitly congruent structure. Therefore inductive inference and heuristic techniques do not explain organisational tendencies because musical structure is understood (and in a certain sense expected) through its implicit structure. Exposure to the peculiarities of a particular system or style does not change the cognition (or expectation) of the implicit structure.

A non-inductive conception of congruence is important because it explains how grammars and schemata can employ different features, with different multiparametrically congruent arrangements, that can be equally as congruent. *GTTM* incorporates a concept of congruence where it is a necessary cognitive tool for the interpretation of structure. From this perspective, similar congruent schemata would emerge in every musical grammar. However, grammars and schemata have varied arrangements of multiparametrically congruent features. Thus, a non-inductive theory of congruence enables a flexible portrayal of differently congruent grammars and schemata.

1.3.2 Aspects of the Tendency for Congruence

The previous subsection establishes that the tendency for congruence in musical structure is a non-inductive product of cognition. This is a top-down constraint that acts on the features of grammars and schemata. It does not act on the perception and cognition of structure in real-time listening, but interacts with cultural constraints during the historical formation of the grammar. Grammars tend towards congruence in a historical period because they were regularly informed by the constraint for congruence in cognition, which explains why parcels of multiparametrically congruent features commonly form in local contexts, described as schemata. As has been pointed out, the tendency for congruence therefore does not explain musical cognition, but merely proposes that congruence is a common product of cognition. A statistical survey of musical structure, such as that which is conducted in the present dissertation (Section 5.2), thus only supports the claim that there is an implicit congruent order in music, and cannot provide insight into how this occurs. Associative-statistical, probabilistic, and inductive models of musical structure, such as the associative-statistical schema theory of Gjerdingen (1988), the 'viewpoints' approach of Conklin and Cleary (1988), Conklin and Witten (1995), and Conklin (2002), or the inductive learning machines of Cope (1996, 2000), are arguably therefore limited as theories of musical cognition (computational models are discussed in Section 3.1). Longuet-Higgins and Steedman (1971, p. 224) argue that music 'would be incomprehensible if the key and metre were called into question before they were established', quoted above. However, since the congruent structuring of music is intuitively cognised, key and metre can therefore be constantly called into question and almost never established, which explains why it is possible to understand highly ambiguous music.

The tendency for congruence occurs in and between the features of different parameters. It is a generalisable notion that analogises across parameters. Terminological descriptions are now presented to describe the various manifestations of congruence. Congruence in a single parameter is termed 'uniparametric congruence' (UC), and between more than two parametric features is described as 'multiparametric congruence' (MC). The tendency for congruence can be formally measured in specific contexts using the rule of multiparametric congruence (the rule of MC). (A more comprehensive explanation of multiparameric congruence is provided in Chapter 4, and models for measuring the uniparametric congruent structure of each parametric feature are presented in Chapter 5.) These concepts are now more fully elaborated.

Uniparametric Congruence

UC is stability or agreement between the elements of a single parametric feature. This can be characterised as simplicity in the structure of a feature, or uniformity between the elements of a feature. For example, UC is tonal or harmonic stability, such as consonant chords in a chord progression, parallelism, or reinforcement between the *Gestalten* of a textural group (the concept of textural grouping is defined in Section 4.2.1), or agreement between the accentuations of beats in a metrical structure. The following uniparametrically congruent features partly define the Classical grammar and butterfly schema (and are quantitatively modelled in Section 5.1):

1. A single tonic key or key area.
2. A regular and hierarchical metrical structure.
3. A chord progression that is close to the tonic in pitch space.
4. A texture that is regular and hierarchical at particular levels.
5. A harmonic rhythm that is regular.

Multiparametric Congruence

While UC describes the stability between the elements of a single feature, MC describes the simultaneous occurrence of more than two uniparametrically congruent features. (Biparametric congruence refers to the simultaneous interaction of two uniparametrically congruent features.) As discussed, congruence is a generalisable and analogical notion, and so applies to a collection of features without explaining the specific interaction between those features. This overall measure of congruence combines the congruence of those parametric features. MC in the butterfly schema and the Classical grammar is the simultaneous occurrence of more than two of the above uniparametrically congruent features (shown in points 1–5 above) (i.e., chords that are close to the tonic in pitch space, regular and hierarchical textural grouping and metrical structure, and simple harmonic rhythms occur together within a single key and regular metrical structure). Uniparametric and multiparametric *noncongruence* also occurs in grammars (elucidated in Sections 5.1.1–5.1.3), but is proposed to occur less frequently because of the tendency for congruence in cognition. Multiparametric noncongruence occurs when unstable features occur together or when unstable or stable features occur together, such as when irregular textures or metrical structures coincide with dissonant harmony, or when regular or irregular harmonic rhythms correspond with irregular textural groups. While congruence presumably constrains all grammars, each can potentially form with different types of implicit multiparametrically congruent structure and with varying degrees of MC.

The Rule of MC

The central claim of this thesis – that a tendency towards congruence in and between the parametric features of the Classical grammar produces butterfly schemata – can now be presented in quantitative terms. The 'rule of MC' is an expression of the tendency for congruence across parameters in particular contexts. The tendency for UC in grammars is similarly formalised using 'the rule of UC'. (The rule of MC actually encompasses the rule of UC since the rule of MC presupposes that uniparametrically congruent features tend to correspond more often than a random collection of uniparametrially congruent and noncongruent features.) The rule of MC is evinced in a musical grammar when particular multiparametrically congruent relationships form more commonly than multiparametrically noncongruent relationships. The Classical grammar and butterfly schema can be defined through particular multiparametrically congruent structures. In the Classical grammar and butterfly schema, the rule of MC would be observed in the common simultaneous combination of the uniparametrically congruent features shown in points 1–5 above. (Evidence that the rule of MC explains the common occurrence of butterfly schemata is provided in a survey of European instrumental music, presented in Section 5.2.)

The rule of MC seems to be a consequence of the cognitive tendency to organise phenomena in the world in the simplest possible way. It is reminiscent of the more widely used 'simplicity principle' (Chatter, 1999; Chatter and Vitányi, 2003). However, by contrast to the inductive claims that accompany explanations of the simplicity principle (Chatter, 1999), the rule of MC does not propose to shed light on the process of cognition, but simply claims that MC is a common product of cognition. Bernstein (1973) tacitly supports the claim that understanding congruence is non-inductive through his conception of ambiguity in musical structure. Ambiguity, which arises through perceived conflict in expectation, necessitates an intuitive understanding of structure, since in order for ambiguity to be experienced by listeners it must controvert an established alternative. This cannot be achieved through comparing statistical expectancies because this would require the internalisation of huge quantities of data by which structural norms might be based (and from which expectations could arise). Such a large quantity of knowledge would be an infeasible acquisition (as argued in the 'poverty of the stimulus argument' of Chomksy (1965, pp. 27–30), described above), suggesting that ambiguity is actually internally felt with reference to *a priori* conceptions of congruence and noncongruence.

The tendency for congruence is tacitly incorporated into generative models, such as *GTTM*, *TPS,* and *CBMS*. Congruence is preferred in the system of well-formedness and preference rules of *GTTM* (Bod, 2002). However, generative theories do not incorporate congruence as a non-inductive concept but use it to explain how musical structure is cognised. This is perhaps a more ambitious enterprise than the present theory and can perhaps lead to a misinterpretation of the grammatical structure, where congruence is inflexibly reified. The rule of MC builds on many other theories that indirectly examine congruent interactions between features (e.g., Rameau (1722), Riemann (1893), Cooper and Meyer (1960), Berry (1976), Lester (1986), Swain (1997, 2002), Pople (2004)), although without directly explaining it as a non-inductive propensity.

Applications of the theoretical principles of generative linguistics directly to music, as in *GTTM*, *TPS*, and *CBMS*, perhaps do not provide coherent explanations of musical structure. Unlike language, instrumental music of the Classical period is non-referential and is not a vehicle for external ideas. As discussed above, the grammatical structure of language is underpinned by a single recursive generative system (Chomksy, 1957, 1965, 1975), whereas music has several interacting systems, such as harmonic structure, grouping structure, metrical structure, etc. These are not well-formed in the same sense as language (e.g., Chomsky (1957)), involving graded well-formed interactions. The rule of MC accords with the common-sense notion that musical understanding is achieved more directly than language because it is a universally understandable medium that necessarily uses the intuitive tools of perception and cognition. Rather than having a number of demarcated universal laws and particular rules, as introduced in generative theories (such as *GTTM*, *TPS*, and *CBMS*), the tendency for congruence is a generalisable constraint for all musical grammars. It permits a conception of the fluidly congruent interaction between the domains of music, and provides a synthesis of the universal and particular, the abstract and concrete, and the local and global aspects of structure.

1.4 Modelling the Butterfly Schema through the Rule of MC and Top-down Hierarchical Selection

This section provides an overview of how this dissertation investigates the tendency for congruence and the rule of MC through a multiparametrically congruent model of the butterfly schema. It also presents an outline of the statistical survey and introduces the theory of top-down hierarchical selection (HS).

The main claim, that the particular manifestation of the tendency for congruence in and between the features of the Classical grammar causes butterfly schemata, requires an explanation of the respective uniparametrically and multiparametrically congruent features. These are overviewed here and more fully elaborated in Chapters 4 and 5. Figure 1.2 shows a typical example of a butterfly schema, in the opening of Beethoven's Piano Sonata Op. 2, No. 3, i (1794–1795), bars 1–4. This is not a definitive illustration of a butterfly schema because it is a category that permits a range of structural descriptions. (It is one aim of this dissertation to define the boundaries of this category, and to uncover its significance.)

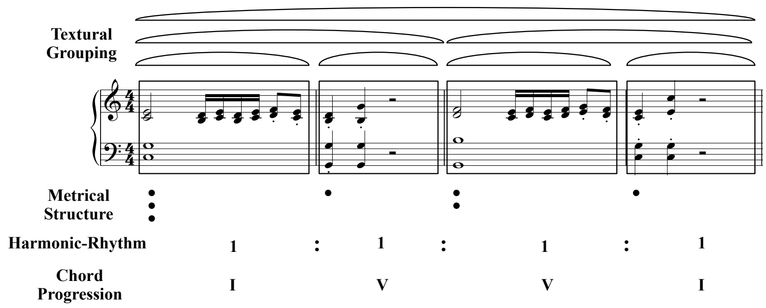


Figure 1.2 A typical multiparametrically congruent butterfly schema in Beethoven Piano Sonata Op. 2 No. 3, i (1794–1795), bars 1–4.

Butterfly schemata have four-part harmonic structures, as shown in the I–V…V–I progression in Figure 1.2. As discussed, butterfly schema progressions have short distances from the tonic in diatonic pitch space (explained using Lerdahl's (2001) theory of pitch space), with a chiastic tension curve that starts and ends on the tonic. The Baroque and Romantic grammars are likewise constrained by the tendency for congruence, but MC is diversely manifest in and between the features of these grammars. Romantic-period grammars often use chords that are greater distances from the tonic in diatonic space than typically used in the Classical period, and also incorporate chromatic spaces (octatonic and hexatonic spaces (Lerdahl, 2001)).

Congruent textural grouping, which is the combined accentuation of voices in texture (and where texture is periodically and hierarchically grouped), occurs at particular levels in grammars, and is distinct in the Classical grammar and butterfly schema. Textural grouping in the Classical grammar informs a particular, but universally constrained, metrical structure that is regular and forms a constituent hierarchy at middle levels of textural and metrical structure — at the level of regular functional harmonic change and at two immediately higher levels. In the Baroque and Romantic grammars, textural grouping informs different metrical structures to those of Classical grammars. The textural grouping and metrical structure of Baroque grammars are more unified at lower levels than in Classical grammars. Romantic grammars, by contrast, have various types of textural grouping and metrical structures. (Section 5.1.2 examines the textures of Baroque and Romantic grammars.)

Simple and regular harmonic rhythm is a distinct and significant feature of the butterfly schema and Classical grammar. Regular harmonic rhythm is ostensibly not a common feature in the grammars of the Baroque and Romantic periods (Section 5.1 covers these divergences). A final model of the butterfly schema, presented in Section 5.1.4, comprises these three main features — harmonic progression, textural grouping, and harmonic rhythm — which are shown to be nuanced structures, connected through their multiparametrically congruent interaction.

It should be noted that merely finding examples of butterfly schemata in a corpus does not constitute evidence that a tendency for congruence constrains the structure of that corpus, because these could be isolated or erratic occurrences. Evidence supporting the tendency for congruence must show that multiparametrically congruent forms are *prevalent* in a distinct historical and geographical corpus. The present paradigm furthers the theory of grammars and schemata because the butterfly schema represents a predominant multiparametrically congruent category that is particular to the Classical instrumental grammar, while also being universally constrained. That is, the butterfly schema is a common product of the particular constraints of the Classical instrumental grammar and the universal constraints of cognition (the tendency for congruence).

Generative approaches, such as *GTTM*, explain schemata inductively, as necessary representations in cognition generated by a listener experienced in a certain idiom. However, schemata form more specifically in many corpuses than pertained in the generalised portrayals in generative theories (such as the 'basic structure' and 'normative structure' in *GTTM*). In the present framework, schemata (as musical structure) are merely influenced by cognition, rather than being strictly caused by it (internally or externally), as will be shown by the flexible depictions attributed to the tendency of congruence. Indeed, a non-inductive framework is essential precisely because schemata are not directly inferred through *a priori* cognitive heuristics — and also because they are contingent on culture. Associative-statistical approaches (such as Gjerdingen, 1988) likewise cannot provide a comprehensive explanation of grammars and schemata, since while they can chart particular distributions, they do not account for the effect of universal cognitive constraints. The present project presents a solution towards greater understanding of musical grammars and schemata, combining generative and associative-statistical approaches, and synthesising rationalist and empiricist metaphysics to provide a flexible approach for modelling grammars and schemata.

That the manifestation of the tendency for congruence in the Classical grammar causes butterfly schemata requires substantial empirical support. A survey was carried out comparing the distribution of the butterfly schema in samples of Classical (*c*. 1750–1800) and Romantic (*c*. 1800–1850) music. Section 5.2 reports on this survey, demonstrating that this particular multiparametrically congruent form is more common in Classical-period samples than Romantic-period samples. Since grammars are variably congruent but tend towards MC, the prevalence of congruent butterfly schemata in the Classical sample over the Romantic sample indicates that the tendency for congruence in this particular form in the Classical grammar causes butterfly schemata. Moreover, since butterfly schemata have features that vary between minimal and maximal congruence, and maximally congruent features are more common than minimally congruent features, then this also suggests that the tendency for congruence causes butterfly schemata.

It is also necessary to show the cultural influence on the Classical grammar and butterfly schema. While all grammars are constrained by the tendency for congruence, the structure of the features and their multiparamerically congruent interaction also depend on cultural selection. The selective action of culture is demonstrated through a top-down HS model, where features are cumulatively selected against each other, revealing how features are formed and uncovering their particular order of selection (presented in Section 5.3). Thus the top-down HS model explains the cultural and historical causes of grammatical and schematic structure. In doing so, this illuminates why features and their order of dependency are particular to grammars. In summary, the aims and objectives of this dissertation are as follows:

Aims

1. To demonstrate that the butterfly schema is caused by a tendency for congruence (measured through the rule of MC) which is manifest in and between the particular features of the Classical grammar.

2. To show that the features of the butterfly schema and Classical grammar are formed through the mechanism of top-down hierarchical selection (HS).

Objectives

1. To show that the features of the Classical grammar and butterfly schema are abstract and explicable in terms of their multiparametrically congruent interaction. (Chapter 4)

2. To formalise the uniparametrically congruent features of the butterfly schema (the chord progression, textural grouping, and harmonic-rhythm ratio) and combine them into a multiparametric model. (Section 5.1)

3. To conduct a statistical analysis of European instrumental music to show that butterfly schemata are more common in the Classical period than the Romantic period, and form with more highly congruent features than noncongruent features. This suggests that butterfly schemata are caused by the particular manifestation of the tendency for congruence in the Classical grammar. (Section 5.2)

4. To illustrate how the features of the butterfly schema and the Classical grammar are selected against each other in a top-down HS model. (Section 5.3)

1.5 Dissertation Structure

Chapter 2, Associative-Statistical and Generative Theories of Musical Structure, examines the depiction of local multiparametric schemata through a review of the literature of associative-statistical and generative theories in music theory. Theories that examine networks of pitch or thematic structures in music, such as Reti (1951) and Keller (1955), propose that these phenomena have a degree of autonomy from the grammatical structure. This is argued to be a limited approach for understanding musical structure because it isolates some musical parameters while neglecting others, resulting in incoherent analyses. The associative-statistical theory of schemata in Gjerdingen (1988, 2007) raises a number of epistemological and ontological issues for schema validation and analysis. Voice-leading schemata are not explicable in terms of their implicitly congruent interrelationships, which is argued to be a limited perspective for understanding the causes of musical structure. A review of generative and proto-generative theories of musical structure, such as Schenker (1935/1979), Cooper and Meyer (1960), and Lerdahl and Jackendoff (1983), shows how schemata can be defined through their generative structural relationships. Schenker (1935/1979) presents the Ursatz ('fundamental structure'), formed from the harmonic series, which is presumed to be the kernel structure from which many pieces of Western music are constructed. Cooper and Meyer (1960) expound fundamental issues that impact on generative theory, such as the congruent relationships between rhythm and metre, and the universal constraints of psychology, including Gestalt principles and the propensity for duple structuring. Lerdahl and Jackendoff (1983) and Lerdahl (2001) present schemata as stable structures in the well-formedness and preference rules of grammar, but are shown to be limited as theories of particular grammars since they present a picture of tonal grammar that is too generalised. They portray cognitive universals as essential, invariant cognitive rules rather than as general principles of cognition.

Chapter 3, Computational, Psychological, and Cultural Theories of Musical Structure, reviews the dichotomy between associative-statistical and generative theories of schemata as demonstrated through computer musicology, psychology, and cultural theories. An exploration of this literature shows how they tacitly provide support for the theory of the rule of MC and top-down HS. Computational musicology presents models of musical structure that express the possibilities of formal analysis, showing the formal uniparametrically congruent implicit structures in music, such as chord progression and metrical structure. Computational models are sometimes limited approaches for explaining how the implicit congruent structure is intuitively understood in cognition. Since the rule of MC is probably the *raison d'etre* of musical structure, models that do not consider multiparametric relationships cannot reveal its implicit congruent and noncongruent structure. However, models that are multiparametric are difficult to implement (e.g., Hamanaka et al. (2005, 2006)). There is no completely satisfactory way to implement multiparametric models because understanding congruence requires rational humans that are intuitively aware of the abstract and implicit structure. Notwithstanding, computational models, such as Tenney and Polansky (1980), Hamanaka et al. (2005), and Marsden (2001, 2005, 2012), present multiparametric analyses that recognise preconceived representations of multiparametrically congruent relationships, although some require various degrees of human input to calibrate this interaction in specific contexts. Inductive and associative models of musical structure, such as Cope (1996, 2000), which recombine abstracted musical signatures of composers to produce compositions in those styles, or 'viewpoints' theories, such as (Conklin, 2002), which formalise associative network relationships, are less powerful since they do not provide generalisable models of musical structure.

A review of foundational psychological literature uncovers limitations in both generative and associative-statistical approaches for modelling schemata. Gestalt psychology and auditory stream segmentation, used widely in generative theories, are often implemented as immutable laws, rather than perceptual constraints, which limits the cultural influence on musical structure. It is suggested that the form of schema theory and prototype theory incorporated in Gjerdingen (1988) is a less illuminating type than presented in many domains of psychology. The versions of schema theory in Rumelhart and Ortony (1977, p. 40) and Rummelhart

(1980, pp. 33–40), show schemata to be operational theories used for procedural tasks in the environment. By contrast, the schema theory of Gjerdingen (1988) does not incorporate the schema concept beyond describing statistical patterns that are abstracted from culture. This is arguably a weaker form of schema theory because it does explain how and why schemata form in cognition or in the world. Also, many prototype theorists argue that the formation of prototypes (or asymmetries) in a category is partly because a category is not just a product of culture, but is constrained by embodied cognition (Rosch, 1975a, 1975b; Rosch and Mervis, 1975; Lakoff, 1987, p. 43). That is, prototype theorists (e.g., Rosch, 1978; Lakoff, 1987) suggest that schemata are a synthesis of universal and particular constraints. This contradicts the culturally situated schemata in Gjerdingen (1988, 2007), which offers no explanation of how schemata might be constrained by universals of cognition.

Behavioural and memetic theories are examined because they provide a basis for the explication of the cultural selection of features in the butterfly schema and Classical grammar. Theories of style hierarchies and causal and selection hierarchies are explored in Meyer (1989), Jan (2001, 2007, 2013), Ellis (2005, 2008), and Okasha (2006, 2012). Selection and causal hierarchies show how phenomena interact in hierarchical relationships through bottom-up and top-down selection. From these theories a notion of hierarchical selection, or causation, is developed, which shows how features in musical hierarchies are dependent on each other.

Chapter 4, Multiparametric Congruence in Butterfly Schemata, formalises schemata as products of the tendency for congruence. Tonality and metrical structure are highly abstract concepts that are variously instantiated in grammars. In real-time listening these highly abstract structures depend on the interaction of the less abstract features of chord progression, textural grouping, and harmonic rhythm. However, when considering the broader cause of grammars and schemata, culturally transmitted forms of metrical and tonal structure condition chord progression, textural grouping, and harmonic rhythm. In this chapter, the levels of abstraction are established where features interact congruently in the Classical grammar and butterfly schema. The implicit multiparametrically congruent interaction between these features is analysed in schemata of the Classical period, and a preliminary version of the butterfly schema is developed. A preliminary version is necessary to explain multiparametrically congruent interaction between features prior to a quantification of the uniparametrically congruent features in Section 5.1. The preliminary version of the butterfly schema comprises the congruent features of chord progression, textural grouping, harmonic-rhythm ratio, and congruent metrical and tonal structure.

The butterfly schema is argued to be a parent schema of many of the voice-leading schemata of Gjerdingen (1988, 2007). A comprehensive critique of the theory, method, and analyses of Gjerdingen (1988, 2007) exposes limitations in his approach for defining schemata. An approach based on MC is required for an explanation and definition of the butterfly schema. The prototypes of voice-leading schemata in Gjerdingen (1988, 2007) are intrinsically multiparametrically congruent, but this implicit congruent structure is not a central conception in the methodology of schema theory. The prototypes of voice-leading schemata are here proposed to be explicable as multiparametrically congruent child schemata of the parent butterfly schema. Many of the examples of the 1–7…4–3 schema identified in Gjerdingen (1988) are incompatible with the rule of MC, since they are not grammatically congruent butterfly schemata. Also, due to the prioritisation of voice-leading structures in validation, many congruent butterfly schemata are not statistically counted in the survey of Gjerdingen (1988). The parallel embedding of schemata in Gjerdingen (2007), which combines grammatically diverse structures, is incompatible with the present theory because the contrasting forms of grammatical congruence in these schemata cannot be simultaneously formed by the rule of MC. It is posited that the focus on associative-statistical relationships and perceptual validity of schemata in Gjerdingen (1988) does not provide a systematic framework for categorising schemata. The butterfly model and the rule of MC provide a system for explaining the implicitly congruent structure of schemata, but require a quantitative measure of the implicit congruence of each feature (presented in Section 5.1).

Chapter 5, Modelling the Butterfly Schema, examines and compares the features of the Baroque, Classical, and Romantic grammars, showing that each grammar has distinct uniparametrically and multiparametrically congruent features. The chord progression, textural grouping, and harmonic rhythm of the butterfly schema and Classical grammar are quantitatively defined, and combined in a multiparametrically congruent model. These features are shown to contrast with the multiparametrically congruent features formed in the Baroque and Romantic grammars.

This chapter reports on a survey of European instrumental music (c. 1750–c. 1850), which compares the distribution of butterfly schemata over the Classical period (c. 1750–c. 1800) and Romantic period (c. 1800–c. 1850). This supports the theory that the rule of MC is manifest in and between the features of the Classical instrumental grammar and produces butterfly schema. The tendency for congruence generates a greater quantity of butterfly schemata in the Classical instrumental grammar than in the Romantic instrumental grammar. Furthermore, the features of the butterfly schema have a larger quantity of maximally congruent features than minimally congruent features, which also shows the influence of the rule of MC. When defined in terms of grammatical congruence, the voice-leading schemata of Gjerdingen (1988, 2007) are shown to be uncommon, meaning the significance of these patterns in the Classical period is questionable.

The final component of the butterfly schema model shows the top-down HS of features in the context of the Classical grammar and universal cognitive constraints (the tendency for congruence). This illustrates how its features follow a particular order of selection, ultimately constrained by the rule of MC (the tendency for congruence). Top-down HS shows the cultural transmission in grammars, and contrast with real-time listening, which works in the opposite direction, which is the bottom-up perceptual-cognitive experience of musical structure.

Chapter 6, Conclusions, provides an appraisal of the tendency for congruence, the multiparametrically congruent model of the butterfly schema, and the model of hierarchical selection. This includes a consideration of the differences between the present theory and generative and associative-statistical theories. A critique of the statistical survey considers various aspects of the methodology, such as the choice of corpora and the significance of the findings. The survey supports the claims of the dissertation, although limitations in the approach are discussed. The potential application of the tendency for congruence in other grammars and schemata is outlined.

Chapter 2: Associative-Statistical and Generative Theories of Musical Structure

This chapter reviews associative-statistical and generative theories of local schemata, which provide a foundation for the tendency for congruence and the multiparametric congruent and top-down hierarchical selection (HS) models of the butterfly schema. Studies that view associative thematic and pitch relationships in the context of a single work or of various corpuses of works are precursors to the later enterprise of schema theory. An examination of Reti (1951) and Keller (1955) in Section 2.1 shows that thematic and pitch theories are limited explanations of musical structure. It is argued that the lack of concern for broader multiparametric relationships means that these approaches do not engage with how musical structure is generated through the wider grammatical structure. A more comprehensive review is provided for seminal schema theories (in Section 2.2), such as those of Meyer (1973, 1989), Gjerdingen (1988, 2007), and Byros (2009a, 2012b). Associative-statistical theories explain schemata through the association of culturally situated features, but they do not explain the multiparameric causes of musical structure, and therefore do no ally with the present notion of the tendency for congruence.

By contrast, generative theories often model local multiparametric structures, such as schemata, as products of stability in the global grammar, tacitly providing support for the tendency for congruence, and are examined in Section 2.3. The tendency for congruence might be viewed as implicit in theories such as Schenker (1935/1979), Lerdahl and Jackendoff (1983), and Lerdahl (2001), although this is not the primary concern of these theories, which mainly focus on how music is cognised. A critique of the relative attributes of associative-statistical theories of schemata, which are specific, and generative theories, which are general, is presented.

2.1 Thematic and Pitch Associations

This section reviews associative theories of thematic and pitch relationships in music. Pertinent theories include Reti (1951), which is concerned with thematic relationships, and Keller (1955), which shows how serial techniques are used in Classical music. Explanations that only consider thematic and pitch relationships are contentious because they do not account for the ubiquitous grammatical features of music, such as harmony, grouping, and metrical structure, and their multiparametric interaction.

2.1.1 Thematic Relationships

Implicit in the theoretical framework of Reti (1951) is the conviction that thematic relationships have a degree of autonomy from the general global grammatical structure (the relationships in and between features, such as as harmony, metrical structure, and phrase grouping). However, themes are necessarily influenced by grammatical structure (multiparametric features), but it is posed to be peripheral to thematic structure in thematic theories. Thematic approaches conflict with many theories of grammars (Schenker, 1935/1979; Lerdahl and Jackendoff, 1983) because the pitch structure of themes is presumed to have little significance outside the context of grammatical structure. Thematic analysis often elects a type of network structure (as explained, for example, in Cohn and Dempster (1992)), because it examines non-hierarchical relationships in pitch. Reti's (1951) approach conflicts with the strict hierarchies and well-formed structures of grammatical approaches, because the latter theories incorporate pitch structure in terms of the grammatical system.

Reti (1951) explores the relationships between themes through the framework of organicism, examining their similarity in the context of a single work. Similarities are recognised through the functions of inversion, retrograde, interversion, and reversion. Due to the subjectivity in the notion of similarity, the methodology is open to criticism. For example, Marsden (2012) argues that similarity between musical ideas is a subjective notion, and therefore it is difficult to base formal models on this conception.

Thematic analysis often requires the isolation of pitch relationships from their harmonic and rhythmic contexts, which conflicts with the tendency for congruence because features in grammars have a necessary (or causal) relationship with each other. Cook (1987b, p. 110) describes the segregation of the parameter of pitch in Reti (1951) as 'unmusical' presumably because of the lack of consideration of multiparametric relationships. Cook (1987b) argues that the 'thematic process' is often illusory because while connections seem to exist between themes these are often unverifiable. The natural brevity of themes and motifs means resemblances are likely to occur that are not of particular significance. Perceived connections between themes or motives can have an association beyond the piece analysed (Cook, 1987b, pp. 109–110). For example, the rhythmic cell of the 'fate' motif from Beethoven's Fifth Symphony is often a generic rhythmic building block for many Classical instrumental works, and so does not explain the particular structure of this work (contrary to the analytical claims in Reti (1951, pp. 165–192)). While distinct categories of themes with exact matches would constitute a basis for a thematic theory, specific likenesses are seldom found in musical structure. Thus the power of Retian analysis to relate themes and explain the structure of works is questionable (Cook, 1987b, p. 110; Bauman, 1952, p. 140).

2.1.2 Serial Analysis of Mozart

Keller (1955) uses serial methods to analyse Classical music. This constitutes a type of associative approach because it gives primacy to the relationships between ordered pitch sets. The analytical technique differs to that of musical set theory, which prioritises unordered pitch sets. The serial analytical method of Keller (1955) uncovers relationships through a process that is distinct from thematic analysis (and also generative approaches). Keller distinguishes between thematicism and serial technique:

> The central point about about serial technique is the unifying function of a certain succession of notes, i.e. a row; and the point about a row is that it need not be a theme or thematic unit, that it is not rhythmically committed. In my

> serial analysis of Strawinsky’s *In Memoriam Dylan Thomas* (*Tempo*, Spring, 1955) I have described a “real” row as “sub-thematic and pre-rhythmic”. (Keller, 1955, p. 13)

In the pre-rhythmic state, tone rows are not musically significant, but are merely abstractions which become meaningful through their context with other features. Keller (1955, p. 14) provides an analysis of Mozart’s String Quartet, K. 156 (1772), i. The four-note row in G major in bars 1–8 of the exposition (marked in numerals in Figure 2.1) is transposed to E minor in the development section, in bars 72–76 (marked in Figure 2.2).

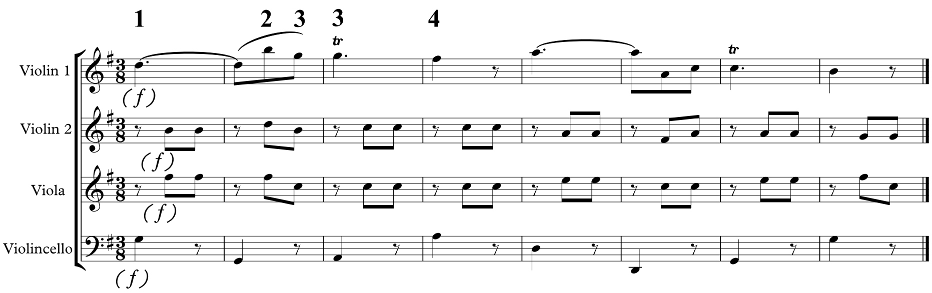


Figure 2.1 Mozart’s String Quartet No. 3 in G Major, K. 156 (1772), i, bars 1–8 (Keller, 1955, p. 14).

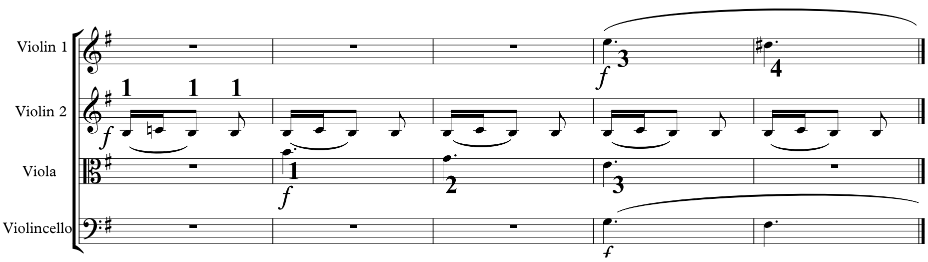


Figure 2.2 Mozart’s String Quartet No. 3 in G Major, K. 156 (1772), i, bars 72–76 (Keller, 1955, p. 14).

Keller (1955) proposes that the relationship between these statements of the row (in Figures 2.1 and 2.2) is distinct from general notions of thematic similarity, because in serial analysis the order of pitch classes must be preserved, whereas in thematic analysis tonal variation between intervallic relationships is permissible (Reti, 1951; Schönberg, 1954). However, a serial approach has limited analytical power for similar reasons to those that have been discussed with respect to Reti’s approach. Moreover, serial analysis is perhaps a more superficial analytical technique than thematic analysis because it only considers pitch order.

Concrete pitch-class relationships are not as generalisable and therefore not as meaningful as explanations that consider the multiparametrically rich relationships that exist in themes.

The primary limitation is that this technique does not consider the interaction between rhythm and pitch, eschewing the significant multiparametrically congruent interactions that are fundamental to the Classical grammar. Many of the features of these examples, such as chord progression and metrical structure, are intrinsically connected through their implicit congruent structure (examined more fully in Chapter 4). Congruent grammatical relationships are more fundamental in the Classical grammar than pitch sets and thematic relationships. The similarity and differences between Figures 2.1 and 2.2 can be seen more clearly through a grammatical analysis. The chord progressions underlying the Mozart rows in Figures 2.1 and 2.2, while different, are both generally uniparametrically congruent, using closely related chords to the tonic (a quantification of the congruent relationships between chords is provided in Section 5.1.1), and are multiparametrically congruent with other features. The rhythm and metre are regular and congruent with the melody, chord progression, and harmonic rhythm, which is typical of the Classical grammar. However, how these are congruent requires further explanation, provided in Chapter 4. Bars 1–8 of Figure 2.1 show a chord progression with a chiastic prolongational tension curve starting and ending on tonic chords (I–V…V–I), a regular and hierarchical texture, and regular and hierarchical metrical structure (constituting a butterfly schema). Figure 2.2 has a different, although not radically different, congruent grammatical structure. The chord progression in bars 72–76 in Figure 2.2 holds the tonic and ends on the dominant (I–V) with broadly congruent metrical and textural structure. These fundamental similarities and differences in the congruent grammatical structures were not perceivable through serial analysis. The exact pitch classes are not significant for a reading of the congruent grammatical structure. The concentration on pitch classes limits the depth of the analytical comparison between the passages. Even in the most generously serial reading of these passages, it is merely peppered with serial elements. The fundamental structural basis of the passages, which shows their underlying similarities and differences, is grammatical congruence.

A major limitation with the approaches of Reti (1951) and Keller (1955) is that they concentrate on a single parameter (pitch class), or group of parameters (thematic structure) without examining the multiparametric interaction between parametric features. In the Classical period, the congruent grammatical structure of broad abstract features is more fundamental than thematic and pitch relationships. Although thematic and pitch relationships are significant, they are rarely verifiable, and pitch sets alone seem to have little significance for explaining the structure of works.

2.2 Changing-note and Voice-leading Theories of Schemata

This section reviews the presentation of changing-note and voice-leading schemata by pertinent schema theorists (e.g., Meyer (1973), Gjerdingen (1988, 2007), Byros (2009a, 2012b)). Broadly, schema theorists view the features of schemata to be connected through associative-statistical relationships situated in particular cultures. Associative-statistical theories of musical schemata contradict many forms of schema theory and prototype theory in other domains of psychology because they do not give primacy to the procedural and functional properties of schemata that are often considered fundamental for a definition of schemata (Rumelhart, 1980). Associative-statistical theories of schemata can often misrepresent the implicit congruent structures of the Clasical grammar through concentration on voice-leading structure. While they show how culturally situated features are present in musical structure, there is no clear system for deciding the preferred network structure of features that constitute schemata.

2.2.1 Changing-note Schemata

Voice-leading schemata have often taken precedence in recent studies (Byros, 2015, pp. 217–220). However, there are various definitions of schemata used in music theory and music psychology. Musical schemata can be any organising structures in music, or any theories of musical knowledge (corresponding with the broad definition of schemata in Rumelhart

(1980) and following the general definition of schemata in systematic musicology, e.g., Leman (1995)). Lerdahl (1991) presents some possibilities:

> There is a wide range of candidates for musical schemata within the classical tonal idiom. One, for instance, is Riemann's (1893) theory of harmonic function, whereby all chords belong to one of three classes. Another is the structure of the 'pitch space' against which listeners evaluate the stability or proximity of events (Lerdahl, 1988). A third concerns the organization of relatively surface patterns of events into types, such as the one extensively studied in Gjerdingen (1988). A fourth encompasses standard tonal forms such as sonata and rondo. (Lerdahl, 1991, p. 273)

This broad range of schema types might render all musical structuring comparable, which perhaps weakens a conceptual basis for schema theory. Meyer (1973) and Meyer and Rosner (1980) provide the first presentation of voice-leading or contour approaches to schema theory (although more often using the terms 'ideal type' or 'archetype'). Meyer (1989, p. 226) defines schemata as discrete local parcels of multiparametric features, prioritising the changing-note structures in the melody. Meyer (1973) establishes a number of schemata which exhibit particular contours, with gap-fill, axial, and changing-note patterns.

The changing-note schema is a precursor to the voice-leading schemata of Gjerdingen (1988, 2007). However, Meyer's (1973) changing-note schemata are less specific than those of Gjerdingen (1988, 2007). Aside from observing the changing-note contour, there is a choice about which exact pitches are used. That is, the changing-note schema can start on the tonic, mediant, or dominant note of the tonic triad (Meyer, 1989, p. 227), thereafter following the voice-leading path dictated by a I–V…V–I harmonic structure. Meyer permits some freedom of harmonic motion because it is assumed that schemata are not fixed entities but fuzzy categories of structure.[8] Changing-note schemata can potentially apply to tonal music in general, rather than a particular or situated musical system. In contrast to Gjerdingen (1988), Meyer (1989, pp. 226–241) presents changing-note schema as part of the 'tonal syntax' common to many subcultures of tonal music. Thus the changing-note schema is common to

---

[8] Kaiser (2007a, p. 76) provides a compelling critique on the demarcation and identification of schemata, especially concerning harmonic features, arguing that identification and validation of schemata is more multifaceted than presented in Gjerdingen (1988). Gjerdingen's approach cannot fully account for aspects such as form function (*Formfunktion*), dimension (*Ausdehnung*), and arrangement (*Inszenierungsweise*), which slip through the analytical filter.

both the Classic and Romantic periods, although it is primarily associated with the Classic period. Romantic composers utilise this schema in veiled form (Meyer, 1989, p. 226).

2.2.2 Voice-leading Schemata

Gjerdingen (1988, 2007) avoids the generality of the changing-note concept in favour of particular pitch constellations, detailing the exact voice-leading events in schemata of the Classical and Galant styles. Gjerdingen (1988) formulates the 1–7…4–3 schema, shown in Figure 2.3. This schema is proposed to be operative during most of the eighteenth century, exhibiting a Gaussian curve of popularity and typicality that peaks in the early 1770s (Gjerdingen, 1988, p. 158).

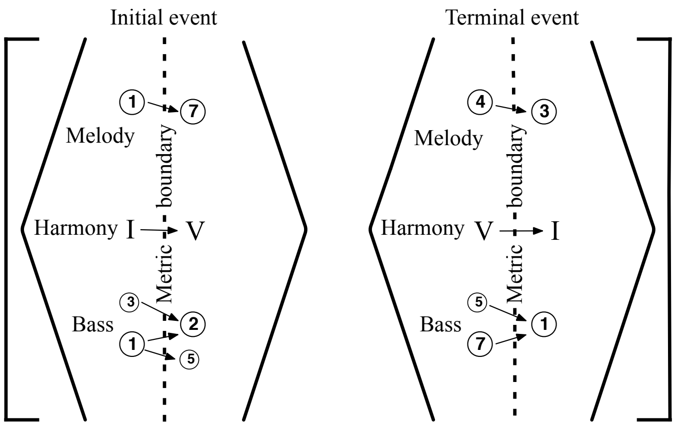


Figure 2.3 Prototype of the 1–7… 4–3 schema (Gjerdingen, 1988, p. 64).

Gjerdingen (2007, p. 21) flatly condemns global theories of musical structure. His schema theory constitutes a microtheory of musical structure (Byros, 2012a), contrasting with the grand narratives of Schenkerian analysis and other global and generative approaches (Gjerdingen, 2007, 21, pp. 434–435). Echoing a similar argument used in postmodern theory (Lyotard, 1979a), grand theories and narratives suffer for being course-grained and therefore generalised, whereas 'little narratives' enable a fine-grained understanding of structure.

While Gjerdingen (1988, p. 158) finds that voice-leading schemata are culturally situated, it is unclear how they relate to the universal or generic patterns of musical structure. Most problematically, Gjerdingen (1988) does not show how voice-leading schemata can be contextualised within the broader universal and generic musical structure. Also, the epistemological status of voice-leading schemata is unclear. Lerdahl (1991, p. 273) questions whether schemata are 'psychologically operative, or […] just figments of our theoretical imaginations?' Since the 1–7…4–3 schema is abstracted by cognition from a continuum of historical change (Gjerdingen, 1988, p. 99), this suggests that voice-leading schemata are primarily cognitive categories. Therefore voice-leading schemata do not have fixed boundaries, but merge seamlessly between one form and another (as such categories are explained in Lakoff (1987), Zbikowski (2002), Neuwirth (2008, p. 402) and Byros (2012a)). This is tenable because musical structure is complex and infinite, but it indicates that in real-time listening, culture is stylistically neutral and is not as significant as perception and cognition for determining schemata, which contradicts a main contention of Gjerdingen's schema theory, that schemata are culturally situated.

It is likely that Gjerdingen considers both mechanisms — the action of culture and cognitive abstraction — to be significant, although a reconciliation of these two constraints is perhaps only vaguely attempted. The mechanism by which perception and cognition abstracts schemata is obscure (Cavett-Dunsby, 1990, p. 84; Lester, 1990, pp. 373–374). Gjerdingen (1988, p. 85) proposes validating schemata (referring primarily to the 1–7…4–3 schema) on the basis of the clarity, distinctness, vividness, and prominence of voice-leading events, but it is unclear how these stipulations could be incorporated into a broader theoretical framework of categorising schemata. Generally, these criteria for a perceptual analysis are not coherently incorporated into an explanation of musical structure.

Cavett-Dunsby (1990), Lester (1990), and Cohn and Dempster (1992), with reference to Gjerdingen (1988), question if there is a basis for preferring one associative network relationship over another. However, while the perceptual-cognitive component of Gjerdingen's (1988) theory of schemata might be problematic, issues surrounding the mental abstraction of schemata are arguably peripheral since the thrust of Gjerdingen's (1988, 2007) theory of schemata is their dependence on culture and history. The statistical survey in Gjerdingen (1988) shows that voice-leading schemata are a product of the association of schema features in a particular time and place. This associationist connection between schema features provides tentative support for the theory of top-down HS of features of grammars and schemata, presented in Section 5.3.

Gjerdingen (1988) builds on the ideas of schema theory presented in Gibson (1969), Schank and Abelson (1977), Rumelhart (1980), and Mandler (1984). Schemata are fuzzy categories formed in cognition around a central prototype, and prototypes are the most salient example of a category. However, in Rosch (1978) and Rumelhart (1980, p. 38) schemata and prototypes are proposed to have a functional, or procedural basis. In the definition of Rummelhart (1980), schemata must be actively used in real-time, as cognitive procedures to operate the world. They are not simply theories about the world. Moreover, Gjerdingen's theory of schemata is markedly different to the schema and prototype theories of Rosch (1978) and Lakoff (1987) in other respects. In the latter prototype theories, categories are formed through the combined interaction of physical, physiological, psychological, and cultural factors. Broadly, they combine embodied cognition with cultural constraints (which is more fully discussed in Section 3.2.2 through an examination of prototype theory). The notion that statistics alone can explain categories is too simplistic, as argued in Chomsky (1957), Rosch (1978), and Lakoff (1987). This accords with the criticism of associative-statistical explanations of schemata in music theory (e.g., Narmour, 1990, p. 30; Lerdahl, 2011). Since Gjerdingen (1988) defines schemata through enumeration, his approach arguably cannot show how schemata are formed through the interaction between cognition and culture. Indeed, the influence of universal cognitive constraints is not explained, which in this dissertation (through the notion of the tendency for congruence) are proposed to be primary influences on the formation of schemata.

The historical distribution of schemata presented in Gjerdingen (1988) arguably cannot explain how listeners, presently or in the past, conceive schemata. A focus on historical distribution does not show how schemata came into existence, are understood, and also abstracted by listeners in real-time. The statistical analysis of the 1–7…4–3 schema in Gjerdingen (1988) is a retrospective and retrodictive venture, which illustrates how a musicologist might examine its history. It might be inferred from Gjerdingen's study that only musicologists are able to consider the expanse of schematic variation through time, and so only they can really understand schemata. This would be a dubious view because it means that schemata are not fully understandable to everyone, except those who have a historical vantage point. This is contrary to a common-sense understanding of music, which is that most people experience music in a broadly similar way. Indeed, the description of music as a universal language is generally warranted (Cook, 1998, p. 127). The concept of 'situated cognition' (Gjerdingen, 2007; Byros, 2009b) is perhaps likewise an unsupportable notion if it contends that certain structures are exclusive, since musical structure is comprehensible from outside cultural spheres. As discussed, since the voice-leading schemata of Gjerdingen (1988) are associative-statistical structures in culture, rather than primarily cognitive representations, Gjerdingen's focus on perception and cognition does not contribute to the cultural situatedness of his theory. Voice-leading schemata are mainly empirically grounded structures in Gjerdingen (1988, 2007), and so it is unclear why perception and cognition should affect these culturally conditioned structures.

The methodology for evincing schemata in Gjerdingen (1988) also requires consideration. In the survey showing the life cycle of the 1–7…4–3 schema, the demarcation of the musical sample is not transparent. Meyer (1989, p. 57) points out the necessity for declaring the limits of samples, otherwise patterns could be cherry-picked to support theories. If we are to accept the distribution of the 1–7…4–3 schema we must know to which data the survey is intended to refer. Limiting the sample by composer and genre is a common approach because this establishes predefined corpora which are arbitrarily demarcated, and therefore less subject to bias (Meyer, 1989, p. 57). In Gjerdingen (1988) there is no information about how certain pieces are chosen for analysis, and why certain genres are examined and not others. That is, although Gjerdingen has explicitly limited the historical window from which the data are extracted (i.e., the eighteenth century), the study is not delimited in terms of other variables, such as corpora, genre, or geography. Therefore it is not evident from the study exactly how popular or typical schemata were in the eighteenth century.

2.3 Generative Theories of Musical Structure

This section reviews proto-generative and generative theories that explore multiparametrically congruent relationships in musical structure. There is an appraisal of traditional explanations of grammatical structure, focussing on the work of and Schenker (1935/1979) and Cooper and Meyer (1960). Following this, the portrayal of local schemata is viewed more specifically, with attention given to *GTTM* (Lerdahl and Jackendoff, 1983) and *TPS* (Lerdahl, 2001). It is argued that generative theories, which consider congruent multiparametric grammatical relationships, provide an explanation of the cognitive influences on musical structures such as local schemata. Moreover, they attempt to integrate the universal cognitive influences with the particular cultural constraints on musical structure. However, generative approaches often generalise congruent relationships in the tonal system, and misrepresent universal laws, such as Gestalt psychology and universals of tonal stability, resulting in readings of musical structure that are not completely coherent with the specific grammars that comprise tonal music.

2.3.1 Schenkerian Theory

This section examines the generative principles underpinning the Schenkerian conception, focussing on the grammatically congruent interaction of harmony with other parameters. Riemannian (Riemann, 1893) and neo-Riemannian theories (Lewin, 1982, 1987; Cohn, 1996) offer a comparible hierarchical perspective on harmonic structure for the late-nineteenth century, viewing hierarchical formal functions of triadic relationships that are reducible to voice-leading operations. However, Riemann (1893) does not present a fully-fledged formal theory, and the neo-Riemannian approach is appropriate for late-nineteenth-century music (and similar grammars). By contrast, the Schenkerian approach is more applicable to eighteenth-century music. This subsection argues that adherence to the tendency for congruence is tacit in the generative, hierarchical, and reductive structures of the Schenkerian conception.

Schenker's *Der Freie Satz* (1935/1979) has been described as a proto-generative theory (Lerdahl, 2012). Schenker (1935/1979) posits that many tonal works are reducible to a kernel structure, which he defines as the '*Ursatz*' ('fundamental structure') (Schenker, 1935/1979, p. 4), shown in Figure 2.4. The *Ursatz* is a contrapuntal elaboration of the major triad and 'an outgrowth of the harmonic series' (Schenker, 1935/1979, p. 10). While the harmonic series is often thought to be the basis of harmonic progressions and function, even prior to Schenker (e.g., Rameau, 1722; Helmholtz, 1863/1875), Schenker provides a more refined analytical framework for examining how the implicit fundamental structure constrains musical elaboration ('diminution'). The *Ursatz* comprises a melody (*Urlinie*) and a bass structure (*Bassbrechung*) that are harmonically and contrapuntally unfolded.

> The *background* in music is represented by a contrapuntal structure which I have designated the *fundamental structure*. *Fundamental line* [*Urlinie*] is the name which I have given to the upper voice of the fundamental structure. It unfolds a chord horizontally while the counterpointing lower voice effects an *arpeggiation* [*Bassbrechung*] of this chord through the upper fifth. (Schenker, 1935/1979, p. 4)

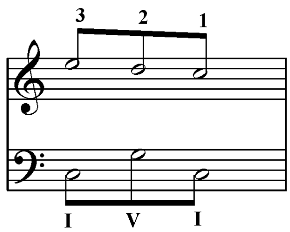


Figure 2.4 The Fundamental Structure (Schenker, 1935/1979, p. 4).

Since tonal music is presumably reducible to the *Ursatz*, this structure might be considered equivalent to the 'deep structure' of generative grammars (e.g., Chomsky (1957)). Consonance and dissonance, which correspond with the tension and relaxation patterns of pieces, are constrained by the *Ursatz*. It prescribes the chord progressions and voice-leading patterns, but also influences other parametric features, such as the rhythmic, textural, and metrical structure. It is implicit on many structural levels, such as foreground, middleground, and background, since surface structure can be reduced to the most abstract global levels (and in practice Schenker does not limit the number of structural levels (Schachter, 1999, 188–189)). An oft-voiced criticism of Schenker's work is that it does not focus on, or give weight to, theories of schematic or thematic structure that can form independently of the hierarchical pitch structure (e.g., Gjerdingen, 2007). While in local contexts the unfolding of harmony might be constrained by voice-leading rules and the adherence to tonic harmony prolongation (Salzer and Schachter, 1969, p. vii), themes and schemata are not necessarily constrained in such a straightforward manner.

The generality of the *Ursatz* as a deep structure of Western musical grammar is debated (e.g., Katz (1945), Salzer (1952), Narmour (1977), Lerdahl and Jackednoff (1983, pp. 290–301), Schachter (1999)). For example, Schachter's (1999) approach is more representative of the conservative group of theorists, proposing that Schenkerian theory be used only to analyse diatonic tonal art music, whereas Katz (1945) and Salzer (1952) are more unconventional, arguing that the principles of Schenkerian theory have application across a variety of musics, such as popular, jazz, pre-tonal, and post-tonal styles (e.g., Mcfarland (2012), Berry (2005)). While the principles of tonal and harmonic prolongation might be universal, the types of prolonged material might not be equivalent in grammars. That is, the foundational diatonic and triadic conception of Schenkerian analysis might be more applicable to musics of certain periods and geographies than others. Schenkerian theory must be modified to suit the various harmonic systems of the world.

Schenkerian theory is intended to provide a theory of tonal music, and although raises questions about universals, it does not provide a systematic method for establishing them. For example, it merely asserts that the *Ursatz* is an *a priori* deep structure and that the major chord is the chord of nature, without providing a theoretical framework based on cognitive principles. The *Ursatz* is limited as a kernel structure since it is not applicable to non-tonal music, only applicable with modification to non-diatonic tonal music (such as octatonic, hexatonic, etc.), and variably applicable to non-Western musical systems. Even in diatonic music, harmonic prolongational patterns of more varied types exist that do not have *Urlinien* and *Bassbrechungen*. This can be seen in the extended tonality of the Romantic period, or in the tenuously related modal structures of Renaissance harmony (as found in the music of Carlo Gesualdo (1566–1613)). Musical styles can be diverse within the tonal realm, and so a demarcation between *Ursatz*- and non-*Ursatz*-based music is difficult to draw. History provides gradually changing musical structures through time, with no essences, only fuzzy categories (Gjerdingen, 1988; Zbikowski, 2002). The designation of a tonal period is problematic because tonality is a generalised construct. Attempting to describe tonal music through its historical context, such as the 'common practice period', is challenging because tonal structure is not easily equated with period and history (Narmour, 1977, p. 136, 1990, p. 20).

While voice-leading rules and tonal and harmonic prolongation are perhaps the most generalisable aspects of Schenkerian theory, applying Schenkerian theory to non-Western musics might be a limited enterprise because harmonic systems are particular to cultural grammars, and voice-leading considerations might not be significant in some musics. For Schenker (1935/1979), voice leading is a universal principle in music. It has been shown theoretically that the principles of toneness that underlie voice-leading rules are universal (Huron, 2001). Although voice-leading rules can be framed as universal constraints that act on musics of various periods and cultures, they are not necessarily manifest in grammars. Narmour (1977, pp. 41–44) contends that a focus on voice leading gives too much focus to strict counterpoint, which might simply not be significant in various contexts.

The Schenkerian approach is limited due to its dependence on human input in the analytical process. Lerdahl (2012) points out that "Schenker was never completely formal" because there is a reliance on the analyst to interpret the unfolding structure (a view shared also by Narmour (1977), among others).[9] The analyst reduces the surface to a middleground structure, and then reduces the middleground structure to the background structure. A related issue is to what extent the *Ursatz* predicts structure and prescribes the unfolding of events or is a *posteriori* description of structure (or a corpus). Narmour frames this problem in terms of the divergence between the top-down action of the *Ursatz* and the bottom-up creativity of the composer:

> All content is determined and eventually assimilated by the *Ursatz*. Historical growth is theoretically accommodated in advance. No work establishes a sufficient degree of originality to absorb and thereby alter the preformed, high-level structure. "Prolongations" never become "fundamental structures." The

[9] Lerdahl views Schenkerian theory in terms of the framework of generative theories, which might not present a fair comparison because these approaches are diverse. Schenkerian theory is used for analysis, whereas generative theories attempt to explain the cognitive interpretation of experienced listeners. *GTTM* and *TPS* might also be described as semi-formal since they deal with conflicting preference rules that require non-algorithmic solutions.

> *Ursatz* remains immune to transformation. The feedback exchange between the accommodating theory and the assimilating analysis works in one direction only. (Narmour, 1977, pp. 41–42)

Narmour (1977, 109) posits here that a claim about the revelatory power of the *Ursatz* is teleological, since the *Ursatz* presupposes what the theory claims to test; that is, it affirms the consequent. Analysts, who are predisposed to look for these structures, might claim that *Ursätze* exist in many works, but whether they are meaningful reductions, rather than assertions about structure, is questionable. Moreover, the variety of grammars in the world would require that differentiated *Ursätze* be used to reduce the particular structuring of those grammars.

Most problematically, rhythm is not systematically treated in Schenker (Narmour, 1977; Lester, 1986). Schenker generally portrays rhythmic features in terms of the connection with the dominant framework of pitch (voice-leading structure and harmonic structure that is constrained by the *Ursatz*). Indeed, Schenker provides few ways to systematically account for situations where rhythmic structure might be an antecedent to pitch structure (Narmour, 1977, pp. 41–42) (the relationship between pitch-led or rhythm-led analysis is problematised in Yeston, 1976, pp. 4–5). This is a significant limitation to uncover the causal seriality of multiparametric relationships, which can differ between particular grammars and pieces. The primacy of harmonic prolongation in Schenker is shown through its priveledged causal influence on other features. A formal account of multiparametric interaction is not provided in Schenkerian theory, but it is often an implicit consideration. Multiparametric rhythmic relationships are also inferred from pitch structure in a semi-formal manner, extrapolated through the 'guiding hand of the analyst' (Schenker, 1935/1979, p. 15).

Analysis is geared towards uncovering causality based on top-down harmonic prolongation constrained by the *Ursatz*, limiting the possible influence of other features in shaping musical structure. The *Ursatz* anticipates *GTTM*'s 'normative structure' (Lerdahl and Jackendoff, 1983, p. 234), which is a similar type of kernel structure that constrains the unfolding of events. However, the normative structure is more generalisable than the *Ursatz* in the realm of tonal music, and is also given context in terms of universal constraints. In *GTTM*, structural freedom is also constrained through the components of time-span reduction and prolongation reduction, which are top-down rules that limit the influence of bottom-up structuring. Everett (2005), Narmour (1977, pp. 168–169), and Lakoff (1987) question the centrality of generative grammatical approaches (in both language and music) because of their incapacity to account for emerging cultural phenomena. Narmour argues that,

> [t]o the critics of transformational grammar, language is not innate but learned, and the observed agreement between speaker and listener is the result of social interaction, not an a priori assumption, since each speaker has his or her own unique idea of the community's spoken language. (Narmour, 1977, p. 168)

While musical structure is informed by social interaction, the strong contention of generative approaches is that music and language are not primarily communicational capacities (discussed in Section 1.2.2, with reference to the views of Chomksy (1957, 1965)). Languages and musical systems are implicit generative systems in cognition. Music is similar to linguistic generative systems in this respect, however, it is also different because it combines a number of diverse hierarchical systems, such as harmonic structure, grouping structure, and metrical structure. Some of the criticisms of Schenkerian theory, such as its top-down bias, essentialism, generality, and inattention to multiparametric properties, are important criticisms that must be addressed in the multiparametrically congruent model of the butterfly schema.

Prolongation is the theoretical focus of Schenkerian theory, concerning the unfolding of tonal and harmonic structure. Horizontalisation refers to the unfolding of harmony, in local and global contexts. Salzer and Schachter (1969) explain the interaction between voice leading and harmony through the horizontalisation of chords.

> The basic idea of chord prolongation is the elaboration in time of a governing vertical sonority – a chord or an interval. Chord prolongation can be achieved by means of several techniques. The most significant of these techniques is the *horizontalization* of intervals belonging to the prolonged chord. When an interval is horizontalized, its tones unfold against a background determined in the vertical dimension by the governing sonority of which it is a part. Horizontalization, therefore, draws into close interrelation the two dimensions of music, the vertical and the horizontal. […] The tones belonging to the triads, of course, need not appear in immediate succession; they may be connected by passing tones or decorated by neighbors. (Salzer and Schachter, 1969, pp. 144–145)

Since horizontalization does not require prolonged notes and chords to be in immediate succession, the degree of abstraction of harmony can vary in global and local contexts. The coherence between horizontalized elements becomes weaker in more extended global contexts because it is more abstract (Salzer and Schachter, 1969, pp. 127–128; Cook, 1987b). In local contexts there is greater correspondence between surface structures and reduction, whereas in global contexts there is weaker correspondence. Local prolongations are therefore less prolongationally abstract than global prolongations (Narmour, 1977, p. 169) being more concrete elaborations of the *Ursatz*. However, as discussed above, the *Ursatz* is a particular structure that constrains the unfolding of a particular grammar, which might not account for the variety of local structures that emerge in different grammars. A more specific kernel form is required for certain grammars and schemata.

In sum, the approach of Schenker (1935/1979) presents a general hierarchical framework for examining tonal music, broadly corresponding with the generative conception. However, a major limitation with Schenkerian theory is that it does not systematically account for multiparametric relationships in musical structure. Unlike the more widely applicable rules of *GTTM* and *TPS*, Schenkerian theory cannot be universally applied since the *Ursatz* contains elements particular to the diatonic system, and universal elements are not demarcated. As discussed, the grammars of other cultures and systems must be reducible to various other types of prolongational kernel structures. The Schenkerian approach would require modification (as in Katz (1945) for example) to permit particular structuring.

2.3.2 *The Rhythmic Structure of Music*

Cooper and Meyer (1960) consider the notion of congruence between various parameters of music, but mainly explore the biparametric relationships between rhythmic and metrical structures. *GTTM* more systematically incorporates MC between parametric features. Longuet-Higgins and Steedman (1971), who broadly base their notion of congruence on Cooper and Meyer (1960), concede that understanding multiparametric interaction is necessary for modelling metrical structure, although suggest that it can evade formal representation:

> In such melodies the metrical structure must be inferred from the harmonic information, and as yet we have not specific ideas as to how this can be done, though a good musician can do it without the slightest difficulty. (Longuet-Higgins and Steedman, 1971, p. 224)

Cooper and Meyer (1960) note that the congruent interaction between rhythmic structure and metrical structure might provide significant understanding of metrical structure, although they are sceptical about the efficacy of formal models to provide insight into how this occurs:

> This division of function among the several elements of music in which some produce unity and others separation is one of the things which at times make it difficult to know what the dominant grouping is. In any particular case is melodic segregation marked enough to outweigh harmonic similarity and temporal proximity? Do instrumental differences dominate grouping in another case, or does harmonic similarity dominate? And so forth. Again, it is partly the fact that no hard and fast rules can be established to solve this problem of the precedence of variables that makes analysis an art rather than a science. (Cooper and Meyer, 1960, 10)

The concern about formalisation in Cooper and Meyer (1960) perhaps rests on the perceived difficulty of modelling grouping structure, which is often considered the most challenging feature of generative grammars to implement (Temperley, 2001, p. 60; Marsden, 2012). Grouping structure is a subjective notion and a more complex structure to parse than other parameters (Marsden, 2012). However, Lerdahl and Jackendoff (1983), Tenney and Polanski (1985), Cambouropoulus (1998), and Temperley (2001), among others, show that the universal constraints of grouping structure can be characterised using Gestalt principles. It is possible that all features of grammars are equally difficult to formalise when viewed outside of the nominal categories generally found in the analysis of Western music. As discussed, tonal and metrical structure, while appearing on the surface to be neatly segmented into bars or uniform tonal areas, are actually more subtly grouped, being more idiosyncratic and changeable than on first approximation, evading formal analysis. This is shown through the variability in these structures in certain grammars (the variability of tonal and metrical structure is discussed more fully in Sections 4.1.3 and 4.1.4).

Notwithstanding their scepticism of formal methods, Cooper and Meyer (1960) present a number of important biparametric rules to connect rhythmic and metrical structure. They examine the effect of various features (harmony, phrase grouping, agogic stress, metrical structure) on grouping and metrical structure, in lower and higher hierarchical levels.[10] Cooper and Meyer primarily use the notions of congruence and noncongruence to depict the biparametric relationships between rhythmic structure and metre. The concept of 'latent metre' (Cooper and Meyer, 1960, p. 89) is used to show how some aspects of the rhythmic structure can cue a metrical structure which are congruent or noncongruent with other aspects of the rhythmic structure. Understanding the interaction between rhythmic and metrical structure is complex because they are interdependent:

> Which comes first, the chicken or the egg? Does rhythm determine meter, or is it the other way around? The answer depends upon the point from which the process is viewed. For the composer, the rhythmic-melodic organization to be projected determines the meter chosen. For the performer, the meter indicated by the composer limits, though it does not in and of itself determine, the possibilities of grouping. For both composer and performer – as well as for the listener – meter establishes a structured continuum of accents and weak beats which acts as a basis for rhythmic and melodic expectation; that is, becomes a norm in the light of which both the regular and irregular are apprehended and felt. (Cooper and Meyer, 1960, p. 96)

Metrical structure can function as a frame upon which rhythmic structure is manipulated, but rhythmic structures are free to construct diverse metrical frames, which is a type of chicken-and-egg dilemma. In the present dissertation, it is argued that metrical structure, while universally constrained, is culturally transmitted in distinct forms, shown by the predominance of distinct metrical types in certain grammars of particular periods of musical history (see Section 4.1.4). Cooper and Meyer (1960, p. 97) discuss the difficulty of distinguishing between universality and particularly of musical structure. Perceptual-cognitive constraints are difficult to separate from stylistic constraints. For example, the commonality of duple groupings in Western music is a potential cognitive-perceptual universal that shapes musical structure, but also seems to be a product of particular cultural constraints:

> Whether such subjective groupings are determined by the nature of human mental processes or whether they arise out of the listener's stylistic habits and dispositions is a question which requires further study. Probably both play a part. But whatever the reason, it appears that in our culture, at least for the past few centuries, such a series of pulses tends to be mentally organized into a duple meter. Any fair sampling of the metric organizations used in Western music will show a clear preponderance of duple over triple organization. For while meter is

[10] The term 'architectonic', used in Cooper and Meyer (1960), is commensurate with 'hierarchical'.

frequently triple on lower metric levels, it is seldom so on higher ones […]. (Cooper and Meyer, 1960, p. 97)

Duple structuring is perceptually and cognitively easier to process, therefore it is more uniparametrically congruent. Moreover, duple structuring permits greater depth of metrical hierarchies (Lerdahl, 2001, p. 286), and greater depth in connection with other hierarchies, which would presumably permit greater MC. However, as an analysis of many non-Western musical cultures might testify, duple structuring is not universal. Many cultures, subcultures, genres, and musical types have triple grouping at certain levels of the grouping and metrical hierarchy. Broadly, considerations of universal and particular constraints require an analysis of the specific grammatical context.

2.3.3 *A Generative Theory of Tonal Music*

This section provides an overview of *A Generative Theory of Tonal Music* (*GTTM*) (Lerdahl and Jackendoff, 1983) to evaluate the extent to which the tendency for congruence is tacit in its system of grammatical rules.[11] *GTTM* generates the preferred analysis as cognised by a listener experienced in the tonal idiom (Lerdahl and Jackendoff, 1983, 1). Reductions of the surface structure of music reveal a recursive system (or 'Humboltian system' in Humboldt (1836/1999)), which 'makes infinite use of finite means' (Chomsky, 2006, p. 15), and combines universal and context-specific rules. There are four component systems in the theory: grouping structure, metrical structure, time-span reduction, and prolongational reduction. Each forms a constituent hierarchy that interacts with the others.

In *GTTM*, three types of rules are used in the four hierarchical components: well-formedness rules, preference rules, and transitional rules. Well-formedness and preference rules are the most significant for the present study. Well-formedness rules define the legal possibilities of musical structure. Preference rules reveal the congruent and noncongruent interrelationships between component hierarchies and parameters, providing a quasi-formal solution to the problem of predicting a preferred structure from a number of possible alternatives.

[11] Lerdahl (2001), Temperley (2001), and Hamanaka (2006) provide extensions to the generative framework of *GTTM*, discussed in Chapter 3.

Transitional rules are necessary to relate prolongational structure and rhythmic structure when an analysis cannot be derived from the well-formedness and preference rules.

Each component hierarchy (grouping structure, metrical structure, time-span reduction, and prolongational reduction) derives its structure partly from interacting with the other hierarchies, but following a distinct order with feedback loops between them. The grouping structure guides the metrical structure, the grouping and metrical structures guide the time-span reduction, and the grouping structure, metrical structure, and time-span reduction guide the prolongational structure. The overall system is a hierarchy of rules, or grammar, where components and features are interdependent and variable. The grouping and metrical structures represent the bottom-up mechanisms (broadly representing perception) which work in a conversely to the top-down mechanisms of time-span reduction and prolongational reduction (broadly representing cognition and culture) (Marsden (2010, p. 4)). The preference rules are particularly important for depicting the congruent and noncongruent relationships between parameters and components. Although the tendency for congruence is not explicitly argued for in *GTTM*, it is often tacit in the well-formedness rules and preference rules. In Chapters 4 and 5 these congruent and noncongruent relationships are more clearly delineated for the Classical grammar and butterfly schema. *GTTM* shows how the system of universal and tonal rules generates local schemata. Schemata are stable structures that form through the well-formedness and preference rules of the grammar (Lerdahl and Jackendoff, 1983, p. 289), supporting the present claim that the tendency for congruence produces local schemata.

While it is presented as a theory of tonal music, a limitation of *GTTM* is that it does not clearly define the particular corpus it is intended to define. That is, it does not clearly demarcate a particular geographical and historical sample, as is required in a stylistic analysis (argued specifically with reference to the *GTTM* in Kerman (1985, pp. 89–90) and as a more general point in Meyer (1989, p. 57)). While Lerdahl and Jackendoff (1983, pp. 4–5) claim that *GTTM* corresponds with 'Western tonal music' of the common practice period (presumably *c*. 1600–1900), it is unclear exactly to what corpus this refers. This means that despite their claims to the contrary, in this respect, *GTTM* does not accord with the scientific criterion of falsifiability. *GTTM* assumes that tonal music is a distinct, unified system.

However, rather than being a well-demarcated class, tonal music is a category with fuzzy boundaries (Narmour, 1977, 1990; Gjerdingen, 1988; Zbikowski, 2002; Kaiser, 2007b). *GTTM* does not discriminate between the grammars and genres of various geographies and periods that are covered by this category (as discussed in Peel and Slawson (1984, pp. 291–292)).

GTTM conceives the tonal grammar as a type of generic system of Western music, implicit in its use of the description 'tonal idiom'. However, it is also closely constrained by universals, psychoacoustics, and psychology. This contrasts with Meyer's (1956) explanation of 'style systems', which generally emphasise the culturally particular properties of musical structure. Musical structure is a close mixture of universal and particular constraints that are not easily teased apart. The blend of universal and particular constraints depends on the structure of the specific grammar. Such a picture of grammars conflicts with the generic presentation of the tonal idiom in *GTTM*. The analysis of tonal grammars requires a more fine-grained analysis since *GTTM* does not explain the particular cultural rules of many grammars, such as the Classical grammar. By contrast, the multiparametrically congruent butterfly schema model includes the particular features of the Classical grammar and a generalisable framework – the tendency for congruence – which incorporates universals.

2.3.3.1 The Congruence Between Grouping Structure and Metrical Structure

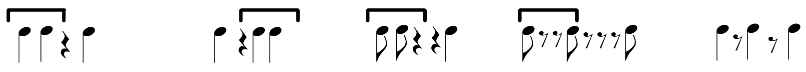
This section assesses how the tendency for congruence is tacit in and between the grouping and metrical structures of *GTTM*. Gestalt psychology, developed in the early-twentieth century, was founded in opposition to the behaviourist movement of the same period. Gestalt psychology proposes that the *a priori* organisation of *Gestalten* is perceptually different to the perception of the individual parts. That is, the whole is qualitatively different to the parts. Perceptual information is organised into groups according to proximity, similarity, closure, symmetry, common fate, continuity, good gestalt, and past experience (Koffka, 1935).[12] The laws of proximity and similarity are most often used in generative theories of music, and can

[12] The law of past experience (Koffka, 1935) might conflict with the associative-statistical approach because this law posits that past experiences influence the immediate perceptual processes.

be viewed as instantiations of the tendency for congruence. In particular, the grouping component of *GTTM* draws from the work of Gestalt psychology (Wertheimer, 1923; Koffka, 1935). GPR 2 (Proximity) (Lerdahl and Jackendoff, 1983, p. 45) follows the Gestalt principle of proximity, and GPR 6 (Parallelism) (Lerdahl and Jackendoff, 1983, p. 51) corresponds with the Gestalt principle of similarity:

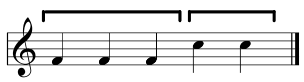
> **Grouping Preference Rule 2 (Proximity)**
> Consider a sequence of four notes $n_1\, n_2\, n_3\, n_4$. All else being equal, the transition n_2–n_3 may be heard as a group boundary if
> a. (Slur/Rest) the interval of time from the end of n_2 to the beginning of n_3 is greater than that from the end of n_1 to the beginning of n_2 and that from the end of n_3 to the beginning of n_4, or if
> b. (Attack-Point) the interval of time between the attack points of n_2 and n_3 is greater than that between the attack points of n_1 and n_2 and that between the attack points of n_3 and n_4.
> (Lerdahl and Jackendoff, 1983, p. 45)
>
> **Grouping Preference Rule 6 (Parallelism)**
> Where two or more segments of the music can be construed as parallel, they preferably form parallel parts of groups.
> (Lerdahl and Jackendoff, 1983, p. 51)

Lerdahl and Jackendoff (1983, pp. 40–41) posit that the principles of perceptual organisation of Gestalt psychology largely constrain melodic grouping structure. Figure 2.5a–e show how proximity determines grouping structure. The first two notes of Figure 2.5a, the last two notes of Figure 2.5b, and the first two notes of Figure 2.5c are strongly grouped according to the Gestalt principle of proximity. However, the first two notes of Figure 2.5d are weakly grouped by proximity. Figure 2.6a–b demonstrates how similarity determines pitch grouping. However, Figure 2.7a–c illustrates the interaction between the Gestalt principles of grouping by proximity and similarity, which results in ambiguity because the preference for similarity conflicts with the preference for proximity. It could be argued, therefore, in contrast to being concrete primitives of a generative theory, Gestalt preferences for proximity and similarity have the potential for conflict, yielding analyses that are frequently ambiguous, suggesting flexible interpretations of structure. Such psychological universals broadly reinforce the notion of the tendency for congruence, since they show how the common implicit structures of music are formed, such as grouping structure and metrical structure. However, flexibilities and idiosyncrasies that actually occur in musical structure show the limitations of formalisation through this analytical system.

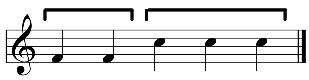


a b c d e

Figure 2.5a–e. Grouping by rhythmic proximity (adapted from Lerdahl and Jackendoff (1983, p. 40)).

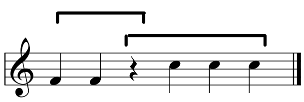


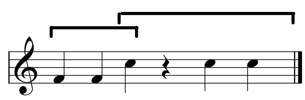
a



b

Figure 2.6a–b. Grouping by pitch similarity (adapted from Lerdahl and Jackendoff (1983, p. 41)).





a b

c

Figure 2.7a–c. Conflict between grouping by similarity and proximity (adapted from Lerdahl and Jackendoff (1983, p. 42)).

The Gestalt principles of proximity and similarity are incorporated into grouping structure through a number of formalisms in *GTTM*, such as 'GPR 3 (Change)' rule (*ibid*., p. 43), which stipulates a grouping boundary when there is accentuation engendered through a change in surface structure. The Gestalt law of similarity is the basis of 'GPR 4 (Intensification)' (*ibid*., p. 49) and 'GPR 5 Symmetry' (*ibid*., pp. 49–50).

In *GTTM*, metrical structure is highly dependent on grouping structure. That is, grouping structure generally cues metrical structure, which means they both are often multiparametrically congruent with each other, according with the tendency for congruence. This manifestation of the tendency for congruence is found in a number of well-formedness rules and preference rules, such as MPR 2 (Strong Beat Early) (*ibid*, p. 76):

> **Metrical Preference Rule 2 (MPR2) (Strong Beat Early)** Weakly prefer a metrical structure in which the strongest beat in a group appears relatively early in the group.

Parallelism in grouping is also posited to provide a cue for metrical structure (Lerdahl and Jackendoff, 1983, p. 75):

> **Metrical Preference Rule 1 (MPR1) (Parallelism)** Where two or more groups or parts of groups can be construed as parallel, they preferably receive parallel metrical structure.

While grouping is primary for the inference of metrical structure in *GTTM,* other types of accentuation are also significant, but to lesser extents. Cues for metrical structure in *GTTM* are found in such factors as note length, texture, bass stability, and the influence of harmony (Lerdahl and Jackendoff, 1983, pp. 76–88). *GTTM* portrays some of these aspects as secondary factors, but they might be of more significance for some grammars. For example, harmony is considered a primary cue for metrical structure by many theorists (Berry, 1976; Lester, 1986; Rothstein, 1995; Mirka, 2009). Indeed, metrical structure can be achieved in a variety of ways depending on the grammar or genre.

Parallelism is difficult to objectify since it is a form of similarity, which is a subjective notion (Marsden, 2012, p. 17). There might be no formal way to model similarity, since it is analogical. Grouping structure, which is based on perceptual similarity, is therefore also a subjective notion. Gestalt laws are too rigid to account for the subtle variances in grammatical structure. By contrast, the tendency for congruence provides a general principle for musical interaction that permits infinite variability of structure.

As discussed, *GTTM* contends that melodic grouping structure is fundamental for the inference of other features (such as metrical structure), yet many writers show that texture is perhaps more significant for cueing multiparametric features (Cone, 1968; Lester, 1986; McKee, 2004). The seriality of the components of *GTTM* is presumed to be universal, in the order described above (i.e., grouping structure, metrical structure, time-span reduction, and prolongational reduction, respectively), but metrical structure can also be formed from texture. (Section 5.1.2 examines how particular textures are congruent with metrical structure.)

2.3.3.2 The Congruence Between Time-Span Reduction and Prolongation Reduction

The intrinsic bias towards a tendency for congruence between the components of grouping structure and metrical structure (and the implicit congruence between the features therein) are likewise found in the relationship between the time-span reduction and prolongational reduction components. Furthermore, all four components generally follow the tendency for congruence when considered as a unified system. An aim of the time-span reduction is to show which pitch events in the surface structure are significant in terms of the bottom-up rhythmically orientated components of grouping and metrical structure (although time-span reduction contains some feedback rules that provide bias for top-down pitch structure) (Lerdahl and Jackendoff, 1983, p. 284). One purpose of the prolongational reduction is to find significant rhythmic elements in terms of dominant top-down pitch prolongations. The prolongational component biases the analysis towards tonic prolongation. This is presumed to represent the listener's cognitive preference for this type of stability in structure (Lerdahl and Jackendoff, 1983, p. 285).

While there is often frequent conflict between the time-span reduction and prolongational reduction, congruence between these systems is nonetheless preferred. That is, multiparametrically congruent pitch and rhythmic structures are favoured in the rule system. However, since the time-span reduction and prolongational reduction can conflict, a transformational rule, the 'interaction principle', is introduced which stipulates the conditions by which the congruence between the influence of rhythm and pitch is established. Lerdahl and Jackendoff describe the interaction principle as follows:

> [T]he Interaction Priniciple states how the time-span reduction regulates prolongational importance. In constructing a prolongational reduction, we can, for each prolongational region, simply search for the strongest prolongational connection possible among the events in the two largest levels of time-span reduction represented in the region. Events of less time-span importance need not be considered. In terms of musical cognition, this means that patterns of tension and relaxation are strongly organized by rhythmic articulation—an intuition that seems obvious, but to our knowledge has not previously been formulated explicitly in the theoretical literature. The power of the Interaction Principle in constraining possible prolongational reductions makes it a central part of our theory. (Lerdahl and Jackendoff, 1983, p. 228)

The interaction principle is necessary to show how the time-span reduction and prolongational reduction interrelate, and is weighted towards rhythmic dominance. Analysis can be pitch-led (prolongational reduction) or rhythm-led (time-span reduction), but both are indicative of the tendency for congruence between these systems. While the biases towards pitch-led or rhythm-led interpretations are explicitly defined in *GTTM*, it is possible that in other grammars the prioritisation of pitch or rhythm dominance is differently achieved.

2.3.3.3 A Generative Definition of Schemata

The theory of schemata in *GTTM* is presented in terms that relate schemata to the global grammar. Lerdahl and Jackendoff posit that schema features ('archetypal features') emerge when well-formedness rules and preference rules of grammar are maximally congruent:

> A passage in which the preference rules maximally reinforce each other, within each component and across components, will be heard as "archetypal". As more conflict appears among preference rules, the passage deviates more from archetypal form and is heard as musically more complex. (Lerdahl and Jackendoff, 1983, p. 289)

In this view, schemata are not independent of a musical grammar, but are stable structures within the grammar. Collections of schema features (such as grouping structure, harmonic progression, and metrical structure) that occur together in local contexts are termed schemata or archetypes.[13] Lerdahl and Jackendoff (1983, p. 289) list the 'archetypal phrase features' formed as a consequence of grammatical stability:

> 1. Each larger-level group is divided into two groups of equal length.
> 2. The larger levels of metrical structure are uniformly duple.

[13] The term 'archetype' is virtually synonymous with 'schema' in this context.

> 3. The grouping and metrical structures are maximally in phase […]. This not only excludes all but the lowest-level upbeats, but also guarantees that the measure lengths of groups are in powers of 2.
> 4. In the time-span reduction, the structural beginning of a phrase is its first event and is a tonic chord in root position; the cadence of the phrase is its last two events and is a tonic perfect cadence.
> 5. The prolongational reduction yields a normative structure […].
> 6. The time-span and prolongational reductions are congruent […].
> (Lerdahl and Jackendoff, 1983, p. 289)

Many of these archetypal features correspond with some of the features of voice-leading schemata in Gjerdingen (1988, 2007), such as the 1–7…4–3, discussed above. Indeed, the 'normative structure' is similar to the features of these voice-leading schemata. It has a distinct prolongational form, which is a tension arc, from low tension, starting on the tonic, to higher tension, moving away from the tonic, and then ending in low tension on the tonic. Many of Gjerdingen's (1988, 2007) voice-leading schemata have a similar tension arc to the normative structure. Lerdahl and Jackendoff describe the normative prolongational structure as the most fundamental structure of tonal music, operating on local and global levels of the tonal system:

> Implicit in the discussion so far is that certain tree patterns recur constantly, whereas others virtually never happen. For example, it is most unlikely that a phrase or piece begins in utmost tension and proceeds more or less uniformly toward relaxation […], or that it begins in relaxation and proceeds toward a conclusion of utmost tension […], or that it begins and ends in tension with a relaxed midpoint […]. These are suggestive possibilities uncharacteristic of Western tonality. Rather, a tonal phrase or piece almost always begins in relative repose, builds toward tension, and relaxes into a resolving cadence […]. This is the most essential way in which the idiom achieves the aesthetic effects of balance and closure. (Lerdahl and Jackendoff, 1983, pp. 197–198)

As discussed, Lerdahl and Jackendoff (1983, p. 289) propose that the stability in the system of tonal grammatical rules indirectly creates schemata. They argue that '[a]rchetypal patterns are not represented directly in the grammar, but emerge as ideally stable structural descriptions produced by the grammar' (Lerdahl & Jackendoff, 1983, p. 289). Lerdahl and Jackendoff (*ibid.*) use the expression 'emerge' to admit a degree of stochasticism in this formalism. Lerdahl (2001, p. 248) concedes that there is an emergent level of analysis in schemata which is distinct from grammatical structure. This suggests that schemata are not completely determined by grammatical rules because the higher-level 'events', the melody and bass structures in the voice-leading schemata (in Gjerdingen, 1988), have combinatorial freedom. However, while schemata might not be completely determined by grammar, Lerdahl (2001, pp. 233–242) demonstrates that they are highly constrained by it.

Whether schemata are determined by grammatical structure or the voice-leading events are significant emergent elements (as proposed in Gjerdingen's (1988) 1–7…4–3 schema) depends on the nature and significance of the voice-leading events. For example, the voice-leading practices of the partimento tradition in early-eighteenth-century Italy is partly assimulated by central European composers later in the century. Musical practice of early-eighteenth century Europe often restricts chord inversions to the 'rule of the octave' (Gjerdingen, 2007, pp. 467–470; Sanguinetti, 2012, pp. 113–125). Focus on scale degree movement in partimento practice is shown in the schematic illustrations of Gjerdingen (1988). While these suggest that voice-leading features might be a significant emergent level of analysis (Lerdahl, 2001, p. 248), it is questionable whether they are separate from the ubiquitous grammatical structure. Many of the voice-leading features operate on higher levels of structure to the grammatical features outlined in *GTTM* and seem to have relatively little influence on those structures.

The incorporation of voice-leading events into models of localised schemata is problematic because during the Classical period chord function is often emancipated from voice-leading structure. That is, chord functions act as independent building blocks of structure (Swain, 2002, pp. 68–82). Thus the chord progressions in schemata contain pitch classes that can be freely mixed and combined, contrary to the explanation of models in Gjerdingen (1988, 2007). (Sections 4.2 and 4.3 show that voice-leading schemata are not essential components in the grammatically congruent model of the butterfly schema). Moreover, in the Classical grammar, voice-leading structures do not form multiparametrically congruent relationships with chord structure, but are merely elements of, and governed by, chord structure. For instance, Komar (1971) demonstrates that voice-leading structures in appoggiaturas and suspensions are often products of the harmonic structure, rather than influences on those structures. Appoggiaturas and suspensions occur when the resolution of dissonant pitch classes is delayed against a harmonic context. This delay of resolution occurs in a particular voice in texture, and so the basic harmonic, rhythmic, and metrical structure remains unchanged by such appoggiaturas and suspensions. However, the model and survey of the butterfly schema (in Sections 5.1 and 5.2, respectively) show that voice-leading events are not essential for modelling grammars and schemata because they are not multiparametrically congruent with the other features of grammars and schemata.

2.3.4 Pitch-Space Theory

Lerdahl's (2001) theory of *Tonal Pitch Space* (*TPS*) mathematically quantifies the figurative distance between pitch, chords, and keys experienced in cognition. The pitch-space theory in *TPS* is supported by Rameau (1722), Riemann (1893), Piston (1941), and Lewin (1987), and also empirical studies (Krumhansl, 1990). Whereas in *GTTM*, little explanation is given to how pitch stability is cognised, Lerdahl (2001) examines the constraints on pitch stability and how these can vary in certain contexts. While musical knowledge can depend on listening experience, musical idioms are largely the product of universal constraints that govern pitch cognition. Pitch space relationships in music are generally based on two attributes: '[p]itches close in log frequency are perceived as near to one another, and so are pitches in a 2:1 frequency ratio' (Lerdahl, 2001, p. 42). These attributes enable a measure of distance between pitch, chords, and keys. There are a number of 'basic spaces' that are the frameworks by which listeners internalise tonal hierarchies. Basic spaces are governed by universal constraints. The basic spaces are diatonic, pentatonic, hexatonic, octatonic, whole tone, and mystic space (Lerdahl, 2001), representing the possible configurations of how pitch distance can be experienced. Certain basic spaces are generally utilised in some grammars but not in others. (Section 5.1.1 quantifies the diatonic pitch-space distances of chord progressions in the butterfly schema.)

The 'constraints on basic spaces' (Lerdahl, 2001, pp. 268–274) have well-formedness conditions and preference rules. The well-formedness conditions of basic spaces are hierarchical organisation, duplication at the octave, and equal adjacent interval classes (Lerdahl, 2001, pp. 268–269). The preferential constraints are larger in number. Firstly, the number of pitch classes at a given hierarchical level of the basic space should approach half those at the next hierarchical level. Secondly, the pitch classes at any level should be as uniformly distributed as possible (Lerdahl, 2001, p. 269). However, Lerdahl (2001, p. 269) notes that '[e]xact halving when combined with maximal evenness is itself undesirable, for it leads to total symmetry, impoverishing [interval class] content and neutralizing position-finding'. 'For a basic space to afford unambiguous position-finding, it must express uniqueness, in which each [pitch class] at a given level has an unduplicated set of intervallic relationships to the other [pitch classes] at that level' (Lerdahl, 2001, pp. 269–270). These constraints are synonymous with the tendency for congruence, showing stability and regularity in pitch organisation. The constraints on basic spaces reduce the cognitive load on memory, permitting complex constructions from comprehensible building blocks. Table 2.1 shows how various scales and chords (which are products of basic spaces) satisfy these constraints.

	sensory consonance	***almost half***	***uniqueness***	***maximal evenness***	***two step sizes***	***two generators***	***small steps***
Scales							
diatonic	yes	yes	yes	yes	yes	yes	yes
pentatonic	yes	yes	yes	yes	yes	yes	no
octatonic	yes	no	no	yes	yes	yes	yes
hexatonic	yes	no	no	no	yes	yes	no
whole-tone	yes	no	no	yes	no	yes	yes
mystic (7-note)	yes	yes	yes	no	yes	no	yes
Chords							
triad (maj/min)	yes	yes	yes	yes	no	no	no
dom7 (Tristan)	yes	yes	yes	yes	no	no	no
aug	yes	no	no	yes	no	yes	no
dim7	yes	no	no	yes	no	yes	no
F6	yes	no	no	no	yes	yes	no
[*0 1 6 7*]	no	no	no	no	yes	yes	no

Table 2.1 Table of basic constraints for scales and chords (Lerdahl, 2001, p. 273).

It can be seen that scales generally conform to the universal constraints of basic spaces. Moreover, basic spaces act as conditioning mechanisms that determine the grammatical and schematic structures that form in grammars. Schemata are generated that cohere with the basic space in use in a grammar. This explains why certain schemata are common in some grammars and virtually non-existent in others. Out of the basic spaces available, the diatonic space satisfies most of the constraints in Table 2.1. The grammar of the Classical period uses diatonic space, which means that the schemata that commonly form should be diatonic.

Since all grammars and schemata are governed by these constraints, often particular grammars and schemata are not unique. Basic spaces can be situated in particular grammars although are not unique to those grammars because they can also be used in other grammars. For example, diatonic space is used in the Classical grammar and also light music of the nineteenth century, and the pentatonic space is often used in many indigenous musics and also impressionist music of Western twentieth-century music.

Lerdahl (2001, pp. 233–242) explains the voice-leading schemata of Gjerdingen (1988, 1996) in terms of chord functions based on pitch-space theory. Figure 2.8 shows the pitch-space functions of Gjerdingen's 1–7…4–3 schema:

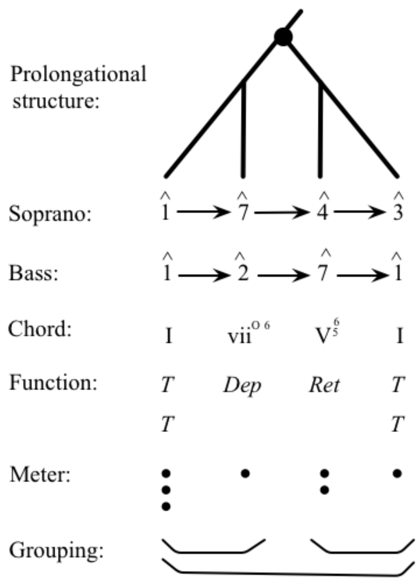


Figure 2.8 Functional model of the 1–7…4–3 schema (Lerdahl, 2001, p. 238).

In Figure 2.8, the harmonic functions, the departure (Dep) and return (Ret), signify a departure and return to the tonic chord, rather than depicting the exact chords. *TPS* conceives schemata as structures that are mainly determined by grammars but which have emergent voice-leading events (Lerdahl, 2001, p. 248). Lerdahl argues that the grammar determines the more fundamental features of schemata, suggesting that schemata,

> [a]re more than surface patterns. They possess grouping, metrical, and tensing-relaxing prolongational structures. These prolongational structures translate into functional sequences that take the form of normative schematic orderings. Listeners attuned to the style usually know where they are in the piece, so to speak, because recognition of a schema also entails awareness of its temporal context. (Lerdahl, 2001, pp. 241–242)

Lerdahl's functional theory of voice-leading schemata shows that their features are fundamentally products of agreement and stability in the grammatical rules, which diminishes the significance of the emergent voice-leading events. However, the claim that the position of a schema in a musical work is prescribed by grammatical structures, such as chord progression, or prolongational tensing-relaxing pattern (Lerdahl, 2001, pp. 241–242), is controversial because it contradicts the significance given to the specificity of the constellation of features which define schemata in Gjerdingen (1988, 2007). Statistical and network explanations of musical schemata, as presented in Gjerdingen (1988), arguably do not explain the formation of schemata in terms of grammatical relationships (Cohn and Dempster, 1992; Temperley, 2001; Lerdahl, 2012).

2.4 Conclusions

This chapter argues that generative grammatical frameworks, such as *GTTM* and *TPS*, are more revelatory about local schemata than associative-statistical frameworks, such as Gjerdingen (1988) and Byros (2009a). As Temperley (2001, pp. 3–4) observes, the ubiquitous structures of grammar – harmony, grouping, and metrical structure – provide a way to explain the occasional structures – local schemata. Generative grammatical frameworks (e.g., Lerdahl and Jackendoff, 1983, p. 289) claim that local schemata form as stable structures of the overall grammar. They therefore provide support for the conception that schemata are a product of the tendency for congruence in grammars.

The tendency for congruence is tacitly incorporated into generative theories and models, but not in associative-statistical theories. The latter, as found in Gjerdingen (1988, 2007), describe network relationships in culturally conditioned schemata. They arguably do not explain how schemata are formed, nor do they frame schemata in terms of the context of universal psychological and psychoacoustic constraints. Gjerdingen (1988) elucidates schemata through observations on their enumerative distribution in connection with their specific voice-leading features. This contrasts with generative theories, which generally concern abstract representations of multiparametric relationships between features (such as harmonic, grouping, and metrical structure).

When local schemata are described in terms of the stable interaction between grammatical rules, as maintained in generative theories, schemata form coherent relations within the global structure. Themes, pitch sets (as discussed in relation to Reti (1951) and Keller (1955) above), and voice-leading schemata (e.g., Gjerdingen (1988)) do not have a grammatical superstructure by which they are connected, being defined by associative relationships that provide no causal explanation about why they exist. A network structure offers no basis by which to prioritise features (as argued in Cavett-Dunsby (1990), Lester (1990), and Cohn and Dempster (1992)). Notwithstanding the advantages of generative theories for conceiving schemata, the discussion in this chapter has suggested that the rules of particular grammars are more fine-grained than presented in the universal and generalised systems of many generative theories. However, the particular features of the grammars of the Baroque, Classical, and Romantic periods require more detailed analysis than given in generative theories. The tendency for congruence builds on the cognitive framework of generative grammars, enabling a more particular model of the butterfly schema (in Chapters 4 and 5). The associative-statistical methodology informs the top-down HS model of the butterfly schema that explains the influence of culture (using the more specific mechanism of selection, explained in Section 5.3). The implicit congruent relationships between features of generative theories and the conditioning relationships between features in associative-statistical theories can thus be synthesised through the tendency for congruence and the theory of top-down HS.

Chapter 3: Computational, Psychological, and Cultural Theories of Musical Structure

This chapter examines the dichotomy of associative-statistical and generative theories in computational musicology, music psychology, behaviourism, and memetics. Computational musicology concerns the formal analysis of musical structure, establishing the quantifiable ubiquitous grammatical features, such as harmonic and metrical structure. It is argued that the tendency for congruence is often a tacit consideration in computational models and analyses. Many computational approaches use associative-statistical or inductive methods to abstract harmonic or metrical structure from music (e.g., Longuet-Higgins and Steedman (1971)). These are problematic because inductive methods cannot explain how these structures are cognised (as argued in Sections 1.2 and 1.3). Also, some computational models generalise idioms for ease of implementation (e.g., Temperley (2001)). This is often the case with the representations of the tonal idiom. Such generalised presentations arguably do not differentiate between the particular types of grammars that constitute the tonal idiom.

While computational analyses show the formal implicit structures of music, they often provide analyses that are merely manifestations of the tendency for congruence (e.g., Povel and Essens (1985) and Rosenthal (1992)). The tendency for congruence can be variably instantiated, and so cannot be encapsulated definitively. Also, since the tendency for congruence is an implicit constraint of computational models, the boundaries between implicit universal cognitive capacities and particular cultural features often evade formal distinction. An argument developed in Section 3.1 is that while computational models indirectly explore the tendency for congruence, this is sometimes inadequately implemented because congruence must be rationally teased apart from cultural features. Computational analysis is unable to account for the variety of structures that are encompassed by the tendency for congruence.

Psychological approaches are also partitioned by the dichotomy of associative-statistical and generative ideologies. In many generative grammars (e.g., Lerdahl and Jackendoff (1983) and Temperley (2001)), Gestalt psychological principles are basically incorporated as fundamental laws of music. However, Gestalt principles should arguably be used as broad constraints rather than strict laws, permitting greater flexibility in the cognitive and cultural influences on musical structure. The use of schema theory and prototype theory in music theory is also reviewed. In some forms of prototype theory and schema theory, in the psychological literature, schemata are not merely descriptive concepts based on empirical observation of the environment, as is generally assumed in the application of schema theory to music (by Gjerdingen (1988, 2007), among others), but are theories of knowledge that show how cognition operates in the environment. Thus schemata do not simply mirror the environment, but permit interpretations of, and action in, the environment. If schema theories are merely descriptive, as in Gjerdingen (1988), they do not explain the cognitive causes of musical structure, and do not justify a framework of schema theory.

Behaviourist and memetic theories support the argument that features of grammars and schemata are selected by culture. Operant conditioning and memetic transmission offer useful mechanisms that explain the cumulative development and evolution of musical structure, which can be applied to the particular features of the Classical grammar and butterfly schema. The memetic theory of cultural transmission provides a way to represent a particular cultural selection history that exists in a certain time and place. Style hierarchies are also examined, which can model the causal interaction of features in musical systems. They show how grammars, governed by universal constraints of psychology, comprise culturally transmitted and selected features. Theories of selection and hierarchy underpin the top-down HS model of the butterfly schema (presented in Section 5.3), which explains the butterfly schema in terms of a broad history of cultural selection and the tendency for congruence.

3.1 Computational Models

This section reviews algorithmic models of musical structure presented in the discipline of computational musicology. Models are examined that present formal descriptions of the grammatical features of music. Models are also reviewed that consider associative-statistical relationships in musical structure (e.g., Bod (2002)). Computational models can provide formal representations of musical features that accord with those of grammatical models in music theory, such as harmonic structure (e.g., Maxwell (1992)) or metrical structure (e.g., Povel and Essens (1985)). Computational analyses that examine associative-statistical relationships, such as the 'signatures' of Cope (1996) or the statistically-induced tonal and metrical structures in Temperley (2007)), are argued to be limited approaches for understanding structural relationships because they do not explain how such structures could actually be cognised.

Many computational analyses (e.g., Winograd (1968), Longuet-Higgins and Steedman (1971), Povel and Essens (1985), and Temperley (2001)) employ pattern-matching systems for parsing musical structures that arguably do not explain the intuitive understanding of these structures in cognition. Multiparametric models can feasibly provide insight into how features commonly interact (as shown in Tenney and Polansky (1980)), but they are limited because they do not show how congruent relationships are recognised with the same coherence and flexibility as is done in cognition. Computational models can be programmed to parse particular relationships between musical features, showing specific manifestations of congruent structuring, similar to such portrayals in generative theories of structure. However, an explanation of the cognition of musical structure cannot be made without showing how cognition intuitively understands multiparametric relationships. At present, cognition and rational understanding remains, to a large extent, a mystery (Chomsky, 2006).

3.1.1 Philosophical Issues in Computational Approaches

Computational approaches can encounter various difficulties in the implementation of generative, associative, and inductive theories. The generative system of rules in *GTTM* is semi-formal, relying on the intuition of the listener to synthesise some of the well-formedness and preference rules in analysis. A pertinent issue for computational implementation of *GTTM* is how to incorporate the preference rules, which interact with each other in ways that require careful calibration of the parameters (Hansen, 2010, pp. 41–42). For example, Hamanaka et al. (2006) calibrate preference rules using automatic and non-automatic adjustable parameters, but are only partially successful because the process by which cognition establishes a preferred structure is elusive, and changes depending on context.

While the implementation of formal models is ostensibly a useful enterprise because it empirically tests cognitive theories of musical structure, computational implementation must overcome the hurdle of explaining rational cognition. Temperley (2001, p. 5) argues that 'while there is no guarantee that such a program performs the process the same way humans do it, such an approach may at least shed some light on the psychological mechanisms involved'. However, although an algorithmic model might illuminate the possible cognitive foundations of musical structure, it might also fail to explain those causes while successfully yielding the output expected by cognition. That is, a model can satisfy the data but not be a feasible explanation of perception and cognition. For instance, some early computational models of key-finding or metre-finding presume to explain the process of cognition through pattern-matching procedures, such as the metrical templates used in Longuet-Higgens and Steedman (1971), or the key-profile approach of Krumhansl and Kessler (1982). Such techniques incorporate predefined models of key structure or harmonic structure by which patterns are successfully identified in musical structure. However, these approaches are not realistic explanations of the cognition of key structure and harmonic structure because they use inductive methods. In cognition, the processes of key-finding or metre-finding are more complexly achieved, involving the intuitive understanding of the multiparametric interaction between features (as discussed in Section 1.3). If the pattern of a system is already known by the analyst, such as the harmonic building blocks of chords or the set of key-profiles of a system, the algorithmic model can be designed to fit these data. However, the process by which these patterns are abstracted in the first place is a more challenging problem. While algorithmic approaches can formally represent key structure and harmonic structure, they arguably do not (attempt to) show how these structures are actually cognised. A related limitation is the generalisation and simplification of structures (such as the abstract representations of harmonic objects and key structure) that are often presented in computational models. This simplification of musical materials is a defensible strategy due to the practical constraints of implementation, but simplification is nonetheless a barrier to understanding.

The difficulties encountered by computational music analysis echo a wider philosophical discussion on the capacity for artificial intelligence to explain cognition. The notion that computers might be able to think and understand as humans do, often associated with 'strong' artificial intelligence, has been vehemently debated by philosophers such as Searle (1980) — who deny strong AI — and others — such as Dennett (1995, 2013) and Chalmers (1995) — who view computational systems as feasible tools for explaining cognition. Searle's (1980) Chinese room argument contends that machines cannot be rational since they merely follow rules and so cannot determine meaning in the world. The argument is as follows. A person locked in a room with no understanding of Chinese is given rules that convert Chinese characters to English characters. The person can perform the task of translating Chinese text into English perfectly without comprehending the meaning of the original Chinese. Moreover, he or she can provide appropriate responses to questions from a person outside the room, appearing to understand Chinese, but without really understanding the meaning of the words and sentences.

Searle argues that an artificially intelligent computer system, provided with knowledge of a grammar and lexicon of a language, likewise does not understand the meaning of the utterances when parsing a text. More significantly, understanding is not possible in a machine since there is no way to ground the real objects in the world to the symbolic language used in computational systems, often termed the 'symbol grounding problem' (Harnad, 1990). This is related to the 'frame problem', which is the notion that artificially intelligent machines have no way of gauging the relevant context that would permit coherent intelligent responses in an emerging and infinite world (McCarthy and Hayes, 1969). Searle's (1980) thought experiment encapsulates one of the major difficulties that hinder research in machine intelligence, that of explaining how rational beings seem to consciously understand the world (which is similar to how Chalmers (1995) frames the ‘hard’ problem of consciousness).

In cognition, harmonic structure is understood in real-time listening through its complex relation with other parametric features, in consciousness, but there does not appear to be a way to construct a machine that is able to do this. Indeed, the fundamental problem with automatic analysis is that machines do not use symbols that are grounded in the real-world environment, and therefore cannot be coherent with the emerging musical structure. The incoherence of automatic analysis is perhaps the root problem of uniparametric and multiparametric models. While multiparametric computational models might enable the examination of the implicit multiparametrically congruent structure of music, they can only partly shed light on these problems. By contrast, the present dissertation does not attempts to solve these mysteries of cognition, but to merely show that the tendency for congruence is a product of cognition.

An important aspect of music theory is to uncover knowledge about how and why musical structure exists (Christensen, 2002, pp. 1–23). While this can be an implicit consideration in computational musicology, it is not always central. Computational musicology has many aims, such as exploring machine intelligence, assisting compositional purposes, developing creative tools, as well as more pedestrian purposes, such as creating publishing software, etc. In computational musicology, it is generally preferred that a model is formal, or that it be computationally implemented, even though these might not be constructive for explaining the rational causes of musical structure. Indeed, formal models and demonstration may impede explanations of musical structure because they present explanations of cognition that are infeasible.

‘Operational’ theories (described in Tenney and Polansky (1980, p. 207)) are of most interest because they seek to establish the cognitive processes influencing musical structure. Such theories are more explicative than ‘descriptive’ theories, which simply mirror the musical structure (although this can also be important when presentations are novel). While operational computational models provide insight into aspects of cognition that can provide specific formal accounts of how structure is arranged, it is questionable whether such models can provide a general account of the cognition of musical structure. Cognition has a way of intuitively grasping the congruent and noncongruent interaction of novel structures in a way that machines cannot yet achieve (as portrayed through the Chinese room argument, and the symbol grounding problem and frame problem by extension). Notwithstanding, computational models, together with linguistic theories, have inspired a number of advancements in formal music theory.[14] Although there are many multiparametric formalisms observed in the following sub-sections, many models consider multiparametrically congruent relationships without directly invoking the concept of the tendency for congruence, excepting those that aim to directly implement preference rules used in generative theories (e.g., Hamanaka et al. (2006)).

Since congruent structures vary in Western grammars, the notion that there are self-contained idioms such as tonal music (as is generally assumed in many computational analyses, e.g., Povel and Essens (1985); Temperley (2001, 2007)) wrongly ascribes an essence to a fuzzy system of features which is associated with various grammars in different times and places. Grammars can only be loosely and arbitrarily subsumed under the category of Western tonal music. In some computational analyses (e.g., Longuet-Higgins and Steedman (1971) and Bod (2002)), corpora are placed under the umbrella of Western tonal music even if they are not prototypical members of this class. Indeed, there is very little evidence that listeners across Western societies hold the same concept for tonal and metrical systems. Meaningful models of the grammars of Western music can only be achieved through fine-grained analysis. Moreover, vague delineations of corpora found in some computational approaches (in many harmonic and metrical models, for example) diminish an examination of the universal and particular forces influencing musical structure.

[14] Computational and linguistic approaches to music were becoming widespread in the decades prior to the publication of *GTTM*. Pertinent examples include Winograd (1968), Longuet-Higgins & Steedman (1971), and Tenney and Polansky (1980).

3.1.2 Key-Finding and Harmonic Analysis Models

As discussed in the previous section, computational analyses of harmonic structure and key-finding algorithms can generalise the varieties of tonal and harmonic structure found in grammars. Some computational models (e.g., Longuet-Higgins and Steedman (1971) and Temperley (2001)) assume that key and harmony are uniform throughout structure and that they do not interact with other parametric features. Computational harmonic analyses can assume a well-formed and recursive hierarchical structure, where harmony is generated from concrete primitive rules, at every scale of structure, from local to global levels. The model of Longuet-Higgins and Steedman (1971) sequentially checks melodies against known scales, eliminating keys through a pattern-matching technique. As discussed above, a limitation of pattern-matching techniques is that they cannot explain how harmonic and key structure is actually cognised because they are explanations based on inductive inference (the problem of inductive inference in the cognition of musical structure is discussed in Section 1.3). Key-finding in cognition requires the intuitive understanding of the implicit multiparametric relationships in and between many musical features. However, it is assumed in Longuet-Higgins and Steedman (1971) that key-finding is simply a matter of matching keys to predefined schematic descriptions, which in this case is achieved through comparing quantity of sharps and flats that are present in a passage.

The harmonic algorithm of Longuet-Higgins and Steedman (1971) is an expert system because it uses *a priori* rules for identifying key structure. It is also brittle because it can only function through the particular input data (the quantity of sharps and flats) and so does not generalise to other musical contexts. Thus the process of key-finding is predefined to suit the solution, again, committing the fallacy of affirming the consequent. That this algorithmic approach is a limited way to understand the inference of key can be demonstrated through a consideration of the actual cognitive process of key-finding. In real-time cognition, the inference of harmony must necessary come before a key can be established. This serial order is obvious because key is an abstraction of harmony over time, in conjunction with the interaction of other musical features (as generally assumed in music theory, e.g., Lerdahl

(2001), Swain (2002). As a first approximation of the process of key-finding, harmony is absorbed in a bottom-up process by listeners in real-time, from which an abstract notion of key is developed and continually updated in the context of the emerging harmonic structure. However, a number of computational models employ the key-finding component separately to the harmonic analysis component (such as Temperley (2001)), which is different to what actually occurs in cognition (where keys are abstractions of the more concrete harmonic structures, as shown in psychological studies, e.g., Krumhansl (1990) and Deutsch (1984)). These computational models simplify and distort these aspects of musical structure.

As discussed, the analysis of key in Longuet-Higgins and Steedman (1971) uses a template for pattern matching and is limited because it is not sensitive to complex multiparametric interaction (between metrical structure, grouping, and texture, for example). However, in broad terms, Longuet-Higgins and Steedman (1971) support the representations of uniparametrically congruent structure (as an isolated parameter), through presenting analyses that generally correspond with the abstract and general notion of key structure. Many other computational models of harmony also present analyses that generally accord with the congruent harmonic structure in generative theories (such as *GTTM*). Winograd (1968) presents a generative algorithm for parsing harmonic structure, with a root-finding component. However, this model likewise does not show multiparametric interactions with other parametric features. Moreover, absent from Winograd's model is a tonic-finding component, which means that the tonic must be inputted by the programmer. This limits its potential as a viable explanation of cognition since, as discussed above, key information is developed from harmonic information in real-time cognition.

Maxwell's (1992) harmonic algorithm is also uniparametric, using predefined harmonic objects to parse chords and their inversions. Again, the algorithm does not contain a tonic-finding component, and so the programmer must input the key. It identifies grammatically congruent analyses of harmonic structure that accord with the depictions of harmonic structure in generative theories, such as *GTTM*. However, similar to the model of Longuet-Higgins and Steedman (1971), Maxwell's model uses an inductive methodology, applying predefined harmonic objects in a top-down process. This pattern-matching of harmonic patterns cannot represent the infinite gradations of harmonic structure in music, nor explain how harmonic structure interacts with other systems (such as metrical structure or phrase structure). Therefore it does not address the more central goals of music theory, which is the exploration of the cognitive, cultural, and universal causes of musical structure.

The Krumhansl and Schmuckler algorithm (Krumhansl, 1990) and the Huron and Parncutt algorithm (Huron and Parncutt, 1993) incorporate key profiles based on psychological probe-tone studies. Probe-tone studies examine the perception of pitch relatedness in the context of a test pitch. These show spatial representations of how listeners relate pitches, chords, and keys. (Listeners broadly relate pitches, chords, and keys similarly across cultures, providing they have sufficient exposure to a musical system (Krumhansl, 1990).) The key profiles are used for the inductive computational analysis of harmonic structure. Similar to the aforementioned criticism, however, inductive methods are problematic because they incorporate *posteriori* expressions of key rather than trying to explain how keys are cognised in real-time. The use of key profiles eschews a consideration of multiparametric interaction, which must be involved in the interpretation of key. Notwithstanding these shortcomings, the Krumhansl and Schmuckler algorithm (Krumhansl, 1990) establishes the structure of chords and keys that broadly cohere with representations in generative theories when considered as isolated parameters.

In an attempt to replicate the neural configurations in cognition that presumably give rise to such aspects as consciousness and rationality, but with a diverging philosophical conception from generative theories, cognitive scientists (e.g., Minsky (1986)) have experimented with connectionist models (or neural networks). Bharucha's (1987, 1991) connectionist harmonic models use networks of nodes to represent pitches, chords, and keys. In Bharucha (1987), pitch nodes activate chords and keys respectively, providing a harmonic analysis that is triggered from the bottom-up association of pitches, showing how chords and keys emerge from concrete note events. In this way, the relationships between pitches, chords, and keys are explicated. A drawback of the model is that it does not integrate a coherent top-down concept of harmonic hierarchy with the bottom-up network structure. The bottom-up network, although feasible, therefore has a bias towards normative pre-conceived harmonic types. Therefore the harmonic objects are asserted as goals of the bottom-up network using an inductive framework, rather than explaining the cognition of these structures.

A limitation with inductive network systems is that it is often not explained why a particular network must be prioritised from a large number of possible alternatives (as discussed in Cohn and Dempster (1992, pp. 171–172)). Bharucha (1987) commits the fallacy of affirming the consequent in his computational model, presenting an algorithm that abstracts a preconceived harmonic structure using a bottom-up connectionist system. By contrast, generative approaches in music theory, such as *GTTM*, examine why such harmonic structures are preferred in cognition, and describe the interaction between bottom-up universal cognitive constraints and top-down particular cultural features. However, *GTTM* does not fully explore the variety of congruence that is manifest in structure (explored in Chapters 4 and 5).

Bayesian probability techniques are used in the key-finding algorithm of Temperley (2007). Bayesian theory permits the structural characteristics of musical features to be interpreted using known information about the relationship between the surface structure and the underlying structure. The key of a passage of music, its structure, is predicted using previous information (sampled from the Essen Folksong Collection in Temperley (2007)) about the relationship between underlying key structure and surface pitch events. A limitation with this type of probability model is that it does not directly interpret the concrete pitch structure that is under analysis. The parsing of structure is based on (prior) statistical information of how sequences of pitches are generically related to keys (Temperley, 2007, p. 56) and so does not adequately account for the actual multiparametric relationships emerging in the concrete structure. Also, it does not account for the top-down cultural constraints that influence multiparametric relationships, and so affect key structure. On pain of repetition, inductive and probability methods cannot explain the cognition of grammatical structure (as argued in Chomsky (1957, p. 17, 1965) with respect to language).

Indeed, probability and statistical methods, such as Bayesian analysis, are often considered *ad hoc* methods for the preliminary testing of hypotheses in some disciplines (as suggested by Martin and Bateson (2007, pp. 143–144) for behavioural analysis), rather than tools for accurate analyses, yet they have a central place in computational musicology and some other humanities disciplines. Chomsky (1957) convincingly argues against the use of statistical analysis for explaining the implicit structure of language because it cannot explain cognition. The dichotomy between inductive and deductive methods of analysis, which forms a central dispute in linguistics (as examined from various perspectives in Klavans and Resnik (1996)), is mirrored in music theory and computational musicology. Viable yet entrenched arguments from both sides of the associative-statistical and generative debate require synthesis. The tendency for congruence is proposed to be such, providing a way to model the universal and particular constraints on musical structure.

The Bayesian key-finding model of Temperley (2007) is biparametric in the sense that pitch features are related to the concept of key. However, important aspects that relate to key, such as rhythmic and metrical features, are not incorporated, as Temperley (2007) acknowledges:

> Some may find the distributional approach of this model (and many other key-finding models) counterintuitive, as it ignores many kinds of important musical knowledge – knowledge about conventional melodic patterns, harmonic progressions and cadences, and so on. To exclude knowledge of this kind from a key-finding model is not, of course, to deny its general importance, but merely to suggest that it may not play much role in key perception. (Temperley, 2007, pp. 96–97)

While some aspects, such as melodic patterns and cadences, might not be crucial to the perception of key (although this is disputable), the forfeit of more fundamental features, such as harmonic, metrical, and grouping hierarchies is palpably detrimental because they are of much greater significance for key-finding. These omissions in Temperley (2007) limit the explanatory power of the key-finding model because they cannot account for multiparametric interaction. Metrical and grouping hierarchies have been argued to be central factors for establishing key in most Western music (as argued in Berry (1976), Lerdahl and Jackendoff (1983), Lester (1986), and Rothstein (1989)). Omitting these relationships means the analysis is less coherent (in addition to the problem of abstraction in probability methods in general). Rational human beings intuitively understand multiparametric relationships and are able to interpret the interaction between features. However, as yet there seems to be no way to determine how this can be done formally in a way which reproduces the flexibility and coherence that is achieved in cognition. Temperley (2007, p. 56) perhaps explains away these important multiparametric features that are necessary for key-finding. The lack of consideration of multiparametric relationships means the interaction between bottom-up (universal and cognitive) and top-down (particular and cultural) structures is also not elucidated in this approach. Notwithstanding these issues, Temperley (2007) provides a system that broadly supports the rule of UC (defined in Section 1.3.2), but within the confines of harmony and key-finding.

An important model that deserves mention in this context is Rohrmeier (2011), which presents a formal generative chord grammar of tonal music (although not computationally implemented). This grammar is a uniparametric approach that strictly adheres to the principles of generative theory (as espoused by Chomsky (1957, 1965)). That is, the model generates harmonic structure from concrete primitive rules. Rohrmeier (2011) thus defines the dependency relationship between functional and hierarchical harmonic structure. The harmonic structure is defined by three functions: pre-dominant, dominant, and tonic. The resulting analyses produce reductions of harmonic structure at various levels of abstraction that describe the harmonic dependencies. This model demonstrates a hierarchical system that incorporates specific generative rules. Similar to other harmonic algorithms examined above, a weakness of the approach is that harmony is treated in isolation, and is generalised.

Indeed, other competing structural hierarchies, such as metrical and grouping structure, also inform harmonic structure. Harmonic reductions are differently conceived and represented when information from other parameters is incorporated. Multiparametric generative theories, such as in *GTTM*, are informal and not fully generative because they do not deal with concrete generative rules (as argued in Rohrmeier and Neuwirth (2015, p. 297)). However, they are advantageous because they show the relationships between multiparametric relationships that are fundamental for cognitive explanations of music. By contrast, Rohrmeier's (2011) model assumes a generic and uniform functional system of harmony that is virtually indifferent to other musical parameters and various cognitive and cultural influences. Notwithstanding, Rohrmeier (2011) provides broad support for uniparametrically congruent grammatical structures such as the butterfly schema, formally demonstrating a functional hierarchical system. (A similar system to define chord progressions is developed in Section 5.1.1.)

Generally, harmonic analysis models implicitly support the rule of UC (although not necessarily the rule of MC (Section 1.3.2 outlines the differences)), because they show the implicit harmonic structure in a single parameter that corresponds with listeners' general representations of abstract chord structure. However, some harmonic analysis models have weaknesses in design because not enough is known about the processes of perception and cognition to understand how this structure is intuitively conceived. Harmonic models cannot interpret the infinite types of multiparametric relationships that form in musical structure. While models establish general uniparametrically congruent structural relationships that are manifestations of the tendency for congruence, they do not explain how these features are actually conceived. A common limitation of harmony and key-finding models therefore is that they do not present a way to account for the real-time bottom-up perception of harmonic data with the top-down abstractions of key. This requires accounting for the interaction between other parameters and differentiating between universal and particular constraints. These problems are parallel to those encountered in the discussions on metrical and grouping models in the following subsections.

3.1.3 Metrical Models

The analysis of metrical structure has received considerable attention in computational musicology. Some metrical models, particularly early examples, such as Longuet-Higgens and Steedman (1971),[15] Steedman (1977), and Longuet-Higgins and Lee (1982), use methods of rhythmic pattern-matching in order to infer metrical structure. Longuet-Higgins and Steedman (1971) present a simple metrical model based on the recognition of the dactyl rhythm (one long accented note followed by two shorter unaccented notes). They argue that this rhythm is necessary to cue metrical structure, following the theory of rhythm and metrical structure in Cooper and Meyer (1960). The significance of the dactyl rhythm for

[15] Longuet-Higgins and Steedman (1971) present two algorithms in their paper. One parses metre, the other harmony.

cueing metrical structure is supported by the Grouping Preference Rule 5 (GPR5) (Symmetry) of *GTTM*: ‘Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length’ (Lerdahl and Jackendoff, 1983, p. 49). The use of the dactyl rhythm for inferring metrical structure is also supported by the congruent interaction between grouping and metre in various other preference rules of *GTTM* (discussed in Section 2.3.4).

Although metrical hierarchical structure is implicit in the dactyl rhythm, this rhythm is not exhaustive of the rhythms found in various metrical structures of grammars. The dactyl rhythm might act as a broad-stroke metrical cue for some Western grammars, but cannot cue metrical structures in grammars where this rhythm is not a foundational structure. While some grammars have regular metres, providing support for the claim that the dactyl rhythm is integral, the dactyl rhythm is of questionable significance for the inference of irregular metres. The notion that metrical structure is inductively inferred through this rhythm is infeasible because cognition cannot use heuristic tools to parse irregular metrical structures. Therefore, this method, although confirming biparametric congruence between rhythm and metre, does not point to a achievable explanation of the cognition of metrical structure.

Longuet-Higgins and Lee (1984), Povel and Essens (1985), and Rosenthal (1992) use the avoidance of syncopation as a criterion for establishing metre. Therefore these models intrinsically view congruence (between rhythmic groups and metre) as an important cognitive consideration. Indeed, the avoidance of syncopation is implicit in the well-formedness and preference rules of *GTTM*. Groups tend to correlate with strong beats in the metrical structure, and accentual events and prolongational structure tend to align with grouping and metrical structure (Lerdahl and Jackendoff, 1983, pp. 75–76, p. 84, p. 220). Povel and Essens (1985) show that the cognition of metre might be achieved with reference to an internal clock. This assumes that metrical structure is fundamentally regular, which although supporting the rule of UC, is a contestable claim because the regularity (or irregularity) of metrical structure depends on the surface structure of the particular grammar and the particular level of hierarchy examined. The Povel and Essens algorithm generates a number of clocks and selects the clock that best matches the input data. This approach is interesting because it provides experimental evidence that the perception of grouping structure changes with metrical context, establishing that conflict can emerge between grouping and metre. This approach shows generally supports the rule of UC because the established metrical structures are generally congruent. However, similar to the aforementioned models, it does not provide insight into multiparametric relationships because it only considers interactions between rhythm and metre.

Metrical projection is explored in Rosenthal (1992),[16] which generates a number of 'ghost events' that are checked against actual event onsets. Rosenthal (1992) improves on Povel and Essens (1985) by generating multiple hierarchical rhythmic levels that are ordered for best fit. However, it is unlikely that simultaneous analyses are generated and prioritised in cognition in this manner. *Ad hoc* trial-and-error processes are unlikely to occur in real-time cognition because it intuitively understands and interprets the 'right' metrical structure through a complex process grounded in the real world (as argued in Sections 1.3 and 3.1.1). An insight of the Rosenthal model (1992, p. 71) is that it shows that metrical structures in music do not always reflect the notated score. Extra layers of metre ('super measures') are generated that correspond with cognitive interpretations, but not with the notated metrical structure (Rosenthal, 1992, p. 72). This is a pertinent insight because it shows the non-essential nature of metrical structure, pointing towards the multiplicity of structural levels above and below the notated bar line that are shown to exist in many grammars (and that exist in the various Western musical grammars, explained in Section 5.1.2). While the Rosenthal model is complex, dynamic, and emergent, it likewise does not account for multiparametric relationships, and thus cannot provide a fine-grained cognitive explanation of metrical structure.

Similar to the Bayesian key-finding algorithm, Temperley (2007) also uses Bayesian techniques to analyse metrical structure. This model infers a structure from a given surface, choosing the most probable metrical grid given prior information about the probability of textural patterns associated with metrical grids. In a sense, it is biparametric because it relates the features of texture and metre, but important features such as harmony and harmonic

[16] The metrical projection model of Rosenthal (1992) anticipates the theory of metrical projection in Hasty (1997), where metrical structure is constructed through real-time projection based on emerging events.

rhythm are not incorporated into the analysis. Temperley (2007, p. 45) acknowledges the absence of many multiparametric features in the model, although the quantity and significance of multiparametric relationships not included in the model is greater than conceded. For example, that '[h]armony is sometimes a factor in meter' (Temperley, 2007, p. 45) greatly undervalues the role of this feature in the Classical grammar, which, as discussed in Section 2.3.3.1, is often regarded as one of the principal cues for metre (noted in Berry (1976), Lester (1986), Rothstein (1995), and Mirka (2009)). In short, multiparametric relationships should be a main consideration for modelling metrical structure. Also, the type of metrical structure that is used in a grammar must be considered, since they are variable, depending on cultural conditioning.

3.1.4 Grouping Models

Grouping models are a challenge to implement because of the subjectivity in delineating grouping boundaries (Temperley, 2001, p. 60) and the elusiveness of multiparametric influences. A number of models often incorporate Gestalt laws of perception to segment musical structure. However, grouping analysis actually requires incorporating information from the disparate domains of rhythm and pitch (although on first approximation this seems counterintuitive because rhythmic features appear to be the primary influence on grouping). Indeed, Tenney and Polansky (1980), Lerdahl and Jackendoff (1983), Bregman (1990), and Gjerdingen (1994) have shown that pitch contour is also crucial for grouping. A problem with some models is that they have not successfully contextualised grouping in terms of the spheres of rhythm and pitch, nor do they incorporate the top-down rules of culture. Also, while grouping is often understood as products of universal processes, such as Gestalt principles, it is also true that they are formed from particular, cultural influences.

Tenney and Polansky's (1980, p. 205) grouping model uses Gestalt principles to segment groups into 'temporal Gestalten'. Using a notion of multidimensional psychological perceptual space, they incorporate the relative influence of pitch contour and proximity to define *Gestalten* (Tenney and Polansky, 1980, p. 211). This cleverly combines the influence of these two parameters in a context-free approach. This is more coherent than calculating absolute values or particular attributes (Tenney and Polansky (1980, p. 211) because it shows the type of biparmetric interaction which occurs in cognition. While the model contextualises two interacting features, it omits significant relationships with other parameters that might also affect grouping. Moreover, it does not admit the influence of culture. Notwithstanding, this model describes the congruent interaction between the disparate domains of rhythm and pitch, providing support for the tendency for congruence. Music theories, such as *GTTM* and *TPS*, have expanded upon this formal exposition of the congruence between these features.

Gjerdingen (1994) presents a contrapuntal grouping analysis model that also employs Gestalt principles. Pitch streams are represented as visual *Gestalten* that track the stream segregation of polyphonic textures. A major drawback with the approach, aside from that the output of the model is difficult to interpret (Temperley, 2001, p. 93), is that it assumes that Gestalt principles are fixed laws of musical structure, rather than broad constraints (critiqued more fully in Section 3.2.1). Temperley's (2001) model likewise presumes Gestalt principles are virtually laws of grouping. However, Temperley's (2001) grouping algorithm brings forth a number of important contributions to the formalisation of grouping structure. Most notable is the Phrase Length Rule (Temperley, 2001, p. 69), which limits the length of groupings to approximately eight notes. The basis of this constraint is psychological: on average phrases last three to five seconds and have roughly five to nine elements due to the limits of short-term memory (Snyder, 2000, p. 13, p. 36). Notwithstanding, this is an *ad hoc* rule which does not provide insight into how listeners actually segment groups; it is merely a heuristic tool that might not be coherent for certain grammars.[17] This rule would not appropriately segment groups in many types of music, such as that of the Baroque (especially J.S. Bach), where long, unbroken phrases are typical in instrumental works.

Temperley's (2001) polyphonic model of texture segments voices into auditory streams based on Gestalt principles, such as proximity, similarity, and good continuation, influenced by the concept of auditory stream segregation espoused in Bregman (1990). Temperley uses a method that consolidates parallel sub-groups of texture into a unified metrical grid using the

[17] Temperley (2001) bases the Phrase Length Rule after an examination of melodies used in Ottman's (1986) *Music for Sight Singing*, which has normative phrase lengths of approximately eight elements. Using this source as a generic benchmark to describe phrase lengths has its limitations, however.

principles of proximity, similarity, and good continuation. This process operates similarly to Lerdahl and Jackendoff's (1983, pp. 153–155) fusion rule, which unifies figurative accompaniment textures into congruent groups that are coherent with the prevailing metre. A similar idea of unifying polyphonic grouping textures to infer metrical structure is outlined informally in *TPS*. In Temperley (2001, p. 100), proximity is the main grouping principle for voices in a stream,[18] incorporated through the Pitch Proximity Rule. The unification of smaller groups into a single larger group assumes that the textural alignment of groups is a preferable and a stylistic constraint in grammars. This is a limited perspective because it conceives contrapuntal structure as necessarily metrical, which might not reflect the various types of contrapuntal structures in grammars. That is, adjacent streams are vertically straightjacketed in texture to permit the inference of a metrical grid, even if the texture might not warrant such alignment. Since the inference of metrical structure is defined by texture, polyphonic texture can conceivably cue irregular metrical structure. Therefore, using this mechanism, polyphonic grouping can never be out of phase with metre. This is counter-intuitive because the elements of texture can be non-isochronous with each other, with voices that are vertically independent. Often voices do not correspond with metre in certain types of polyphonic music of the Renaissance and Baroque (such as that of J.S. Bach).

The non-isochronous nature of some polyphonic grouping appears to present an awkward obstacle for generative theories. Voices in polyphonic textures are often rhythmically noncongruent with each other, being contrarhythmic and heterorhythmic, rather than homorhythmic, as in the Classical grammar. While the tendency for congruence is a universal constraint, in polyphonic music the conflicting voices necessarily diminish the overall congruence in the system. These independent voices can presumably each imply multiparametrically congruent relationships with other features (such as the chord progression and metrical structure), which can create implicit conflict (noncongruence) with the implicated congruent relationships of other voices. That is, the multiparametrically congruent accentuation suggested by one voice in texture can conflict with those suggested by another voice in that texture, creating myriads of conflicting multiparametric

[18] This principle is demonstrated in the pseudo-polyphony on a single-line instrument, where a single melodic thread oscillates between registers creating the illusion of two separate streams. Many of J.S. Bach's works, typically in violin passages, make use of pseudo-polyphony.

relationships.[19] This seems incoherent with the claim of a tendency for congruence, which presumably would imply that for congruence in texture to be the norm, accentuations must correspond. However, while polyphonic voices can be noncongruent at some levels of metrical hierarchies, they generally reinforce each other consistently at other levels. Indeed, distinct types of congruence and noncongruence at particular levels of structure are common in grammars generally (as shown in the texture of the Classical and Baroque grammars in Section 5.1.2). The relatively high noncongruence in polyphonic textures is not considered in the model of Temperley (2001), which basically depicts a constraint for homorhythmicity. These nuances of grammars, which have congruence at certain levels of texture, must be understood to yield an accurate picture of metrical structure in various contexts.

Bod (2001b, 2001c, 2002) provides a model of grouping perception that combines statistical probability with a preference for grammatical simplicity. Bod (2002) incorporates the 'likelihood principle' (based on inductive inference), and the 'simplicity principle', based on the preference for simple structures in cognition (explained in Chatter (1999)). These yield accurate depictions of the phrasing of music and language. Interestingly, Bod (2002) shows that Gestalt perceptual principles are not necessary to parse the grouping structure of the Essen folksong collection:

> More than 32% of the Essen folksongs contain a phrase boundary that falls just *before* or *after* a large pitch or time interval (which we have called *jump-phrases* …) rather than *on* such intervals — as would be predicted by the Gestalt principles. We have also shown that for 98% of these jump-phrases, higher-level phenomena such as melodic parallelism, meter or harmony are not of any help: they simply reinforce the incorrect predictions made by the Gestalt principles […]. (Bod, 2002, p. 14)

That Bod (2001b, 2001c, 2002) finds Gestalt principles are not significant for organising grouping in selections from the Essen Folksong Collection, means the induced internal structures of the melodies yield contrary structures to those that would be ascribed through Gestalt principles. According to Gestalt principles, leaps in melodic structure should be accompanied by grouping boundaries, but Bod shows that leaps often occur before or after

[19] The view that voices in polyphonic music might be rhythmically noncongruent (that is, contrarhythmic, or out of phase with each other) is not controversial in music theory, and has been examined in a number of other theories (Berry, 1976; Benjamin, 1984; Lester, 1986).

the real phrase boundary.[20] However, the criteria for the grouping model in Bod (2002) are obscure, referring to a vague notion of internal structure that is not fully elaborated. A further problem with Bod's models is that they do not make explicit the influence of multiparametric interaction. Specifically, the models do not incorporate metrical and harmonic information because they are presumed not to be significant (Bod, 2002, p. 14).

3.1.5 Multiparametric Models

Multiparametric computational models are potentially more feasible explanations of cognition than uniparametric models because they can account for the interactions between multiparametric features that are necessary to fully understand musical structure. However, multiparametric models are the most challenging to implement. The interaction between features is arguably one of the most intriguing aspects of music theory because it shows the connectedness of features that operate in the disparate domains of rhythm and pitch. Many music theories implicitly and explicitly show the common relationships between multiparametric features (Berry, 1976; Yeston, 1976; Lerdahl and Jackendoff, 1983; Benjamin, 1984; Lester, 1986).

Hamanaka et al. (2005) have provided a partial implementation of *GTTM*, presenting an automatic time-span tree analyser (ATTA). In *GTTM*, preference rules, although designed to tacitly prefer congruence, also permit conflict, and so represent preferred and conflicting analyses. The ATTA numerically weights preference rules and has manual adjustment mechanisms to calibrate the interaction. This approach describes the rich picture of the interaction between multiparametric features in the well-formedness and preference rules of *GTTM*. However, the difficulty of quantifying the interaction between preference rules is not fully overcome. For example, symmetry and parallelism can be construed as types of similarity, which are subjective, presenting a problem for implementation (as argued by Marsden (2012)). Rules that demarcate parallelism in *GTTM*, such as those incorporated into the grouping and metrical components, cannot be formally defined (Hamanaka et al., 2005, p. 359). This broader philosophical issue impeding the modelling of human cognition is foundational to computational musicology (see Hansen (2010) and Marsden (2012)).

---

20 The actual phrase boundaries are marked in the Essen Folksong Collection.

Since preference rules are manually calibrated in the ATTA, the model is of debatable automaticity. Manual adjustments are required to weight preference rule relationships to the specifications of musical structure. Hamanaka et al. (2007) provide further development of the ATTA model, presenting a full automatic time-span tree analyser, the FATTA. Hamanaka et al. (2007) purport that this model automatises adjustments for the construction of a time-span tree. However, it is only partially successful because it actually presents flexible biases for creating time-span reductions. This process is not fully representative of how time-span reductions might be created in cognition, and does not explain how they might be differently formed in other contexts. Since the concept of similarity in grouping and melody is a subjective notion (Marsden, 2009, p. 4; Temperley, 2001, p. 60), the concept of an automatic analyser is in principle problematic. Notwithstanding these criticisms, the theoretical insight of these models is that they present developments towards formalisation of the implicit musical structure. Future work in this area might explore how preference rule conflicts can be adjusted for different grammatical contexts, providing greater understanding of multiparametric relationships.

The multiple viewpoints models of Conklin and Cleary (1988), Conklin and Witten (1995), Conklin (2002), and Pearce and Wiggins (2006, 2012), also show inductive approaches to multiparametric modelling. The model records the multiparametric surface structure, abstracting various viewpoints that can be used to calculate predictive data about musical structure from which to generate new pieces. The predictability of particular combinations of viewpoints is calculated using Shannon's (1951) entropy function. This measures the viewpoints with the lowest entropy that would provide the most predictable forms in the style. Best-fit combinations of viewpoints are the most predictable (See Conklin and Cleary (1988, p. 3), Conklin and Witten (1995, p. 31)). Therefore the viewpoints approach provides support for the generative models and the notion of the tendency for congruence, since optimal viewpoint combinations have low entropy, which are the most congruent. However, viewpoint approaches use inductive methods, and so cannot explain the cognition of the multiparametric interaction of features. Viewpoint approaches are in this sense contrary to the present approach.

Conklin and Witten (1995), using Bach chorale melodies, abstract low-entropy combinations across parameters, calculating predictable stylistic combinations. This yields useful data, such as how scale degree is correlated with an event that begins on the first beat of the bar (Conklin and Witten, 1995, pp. 28–29). This corresponds with the observations of generative theories (e.g., Lerdahl and Jackendoff (1983, p. 165)), and supports the notion that the tendency for congruence is implicit in musical structure. However, the approach is problematic because it is a based on statistical probability. The use of viewpoints rests on a probabilistic cross-entropy calculation that does not provide a direct explanation of the implicit generative structure. Like many probabilistic methods, viewpoints merely describe statistical expectations, but do not explain cognition. Also, explanations about the relationships between bottom-up universals and top-down factors of culture are not provided, such as the norms of top-down harmonic prolongation in certain grammars.

A consideration of multiparametric relationships is tacitly incorporated into the models of Marsden (2007, 2010, 2013), which present generative automated analyses of musical structure. Marsden develops a number of analytical systems that automate Schenkerian analysis and *GTTM*. The multiparametric hierarchical grammatical conception is shared by both Schenker (1935/1979) and *GTTM*. Marsden focuses on Schenkerian theory, pointing out that,

> [w]hile the theory of Lerdahl and Jackendoff [1983] has the advantage of systematic description, it does not, in my view, give a sufficiently detailed description of a musical structure. It describes a structure of melody plus harmonic support rather than a full contrapuntal structure. I therefore choose to base a structural description on Schenkerian theory. (Marsden, 2007, p. 56)

While Marsden argues that a limitation of *GTTM* is its simplistic treatment of contrapuntal structure, the research programme often combines the insights of *GTTM* with Schenkerian analysis, in doing so furthering initial investigations on automated Schenkerian analysis by Smoliar (1980) and Kassler (1977). Marsden (2001) sees musical structure as networks of elaborations that are inherently contrapuntal, and implicitly multiparametric. For each note, or 'place', there is a multiplicity of possible elaborations, termed E-graphs (Marsden, 2001, p. 39). A persistent problem with implementing Schenkerian analysis is the multiplicity of permissible analyses that are generated as potential *Ursätze*, and the ambiguity in the selection of the 'correct' alternative. This problem is consistent with many network models, as argued in Cohn and Dempster (1992), because there is no basis for deciding a preferred structure. Reducing the rules of reduction in a formal Schenkerian analysis requires filtering out an extremely large number of possibilities. Finding the principles by which they should be selected is a challenge (Marsden, 2011, p. 673). Marsden points out that,

> [i]f we temporarily disregard the constraints which make a graph valid in Schenkerian terms and assume that all trees are binary, the number of possible analyses of a piece is at least as many as the number of binary trees possible with *n* terminal nodes, where *n* is the maximum number of nodes in any voice in the piece. This is the 'Catalan number' C_n: (2n)!/(n+1)!n!. Thus we can expect the solution space for Schenkerian analyses of a piece to grow factorially with the size of that piece (Marsden, 2007, pp. 56–57).

Despite the large analytical possibilities, Marsden (2007, pp. 56–57) maintains that a definitive *Ursatz* is tractable through a process of limiting the possible analyses generated. From these manageable computations of *Ursätze* a preferred structure is chosen. In the selection of an *Ursatz*, Marsden (2010) combines bottom-up elaborations with the top-down preferences for *Ursatz* structures. The selection of potential *Ursätze* requires establishing rhythmic and pitch points in structure that are structurally significant. Marsden's models find feasible *Ursätze* that have been partially corroborated by 'experts' (Marsden, 2013). However, since models are concerned with the tonal *Ursatz*, they might only be superficially generalisable to non-tonal and non-Western works. Marsden (2013, pp. 361–362) examines issues of universality and particularity, but his models do not yet provide a satisfactory way to show how these relationships in musical structure are cognised by listeners. Neither do they satisfactorily explain how bottom-up and top-down relationships interact in cognition.

Cope (1996, 2000b) uses inductive methods for generating musical structure, through computer-aided composition, abstracting 'signatures' in compositions that are 'recombined' in new compositions. Using the reverse process to the prioritisation and abstraction of schemata in perception and cognition in Gjerdingen (1988), Cope shows how common signatures can be abstracted. In Cope's (1996, 2000b) compositional machines (EMI and

Alice), signatures are identified and used to create various pieces. In an earlier book, Cope defined signatures as,

> [a] set of contiguous intervals (i.e., exempt from key differences) found in more than one work by the same composer. Signatures typically contain two to nine notes melodically and more if combined with harmonies. They are generally divorced from rhythm, though rhythmic ratios often remain intact. Signatures are work-independent. They do not sound as if they come from a certain work of a composer but rather from the style of a composer (or a group of composers). (Cope, 1991a, p. 46)

Cope (1991a, p. 46) contends that signatures are 'generally divorced from rhythm'. However, this claim is contradicted in Cope (2000b, p. 109), where 'signatures are typically two to five beats (four to ten melodic notes) in length and usually consist of composites of melody, harmony, and rhythm'. It is difficult to ascertain if the grammatical superstructure of Cope's machines are sensitive to the particular characteristics of local signatures, which are segments of musical information, since the automated mechanism of extraction is not clearly separable form human input. Also, since signatures can isolate pitch sequences from key, rhythm, metre, and pitch, and in doing so change their stylistic meaning, there might be incoherence between the top-down and bottom-up processes in this model. Temperley (1995) argues that changing the metrical context of a motivic pattern radically changes its meaning for listeners, suggesting that Cope's method might be limited through its arbitrary abstraction of signatures of a single parameter.

EMI and Alice (Cope, 1996, 2000b) have varying degrees of autonomy from manual input, but the role of this input is somewhat abstruse (Wiggins, 2008, pp. 110–111). The separation between human input (which includes tweaking parameters to act as filters) and the mechanical abstraction of signatures is not clear. There might be many valid insights about machine automation — and also novel explications of style — which cannot be understood because what aspects are the product of machine, and what are the outcomes of human input is not demarcated. Therefore gleaning useful information about the interaction between parametric features is challenging. In the interests of scientific reproducibility, Cope (1996, 2000b) might have provided a full explication of the algorithms used that would permit others to test the claims:

> [I]n all of [Cope's] books, and in the peer-reviewed papers (predominantly in the Computer Music Journal) of which they are mostly composed, I have been unable to find published details (to the extent of reproducibility) of how they work—rather, there are imprecise discussions of representations and rules, filled out with examples that sometimes give us an illusion of understanding what the mechanism does. (Wiggins, 2008, pp. 110–111)

An assumption in Cope (1996, 2000b) is that it is possible to formally differentiate between the styles of composers. However, individual composer styles, while intuitively distinguishable, are much more difficult to distinguish quantitatively than periodic styles. The difference between periodic or historical grammar is often much more marked than composer style, although some composers (e.g., J.S. Bach) have styles that are distinctive and might be quantifiable.

A further problematic aspect of Cope (1996, 2001) is how the relationship between top-down and bottom-up constraints on musical structure is conceived. Cope's models generate signatures that are interspersed in a harmonic structure loosely based on a Schenkerian prolongational framework. However, the procedure for hierarchically arranging signatures within a top-down prolongational superstructure is not fully established. That is, it is not made transparent where the Schenkerian superstructure ends and where the recombined signatures of the particular corpus begin. A Schenkerian system, which is context- and grammar-dependent, might not be suitable for representing some grammars. These details are obscured beneath discussion of a more philosophical and aesthetic type, accompanied with discussion on mechanical matters. It is thus not evident which relationships are universally implicit, which belong to the tonal idiom, and which to the particular composer. There is also a lack of clarity about the trends of multiparametric interaction. In general, it seems there is a conflation between the artistic exploration of algorithmic composition with the scientific goal of analysing musical style. EMI and Alice might have value as tools for music making, and the arguments within his books raise important philosophical ideas about the nature of art, such as exploring the emerging structures of music through algorithmic composition. However, aspects of music theory concerned with the abstraction of signatures, and the relationship between bottom-up and top-down structure, require further development.

3.2 Psychological and Cultural Theories

This section reviews psychological and cultural theories, particularly Gestalt theory, behaviourism, and memetics, which provide a basis for the tendency for congruence and the multiparametrically congruent and top-down hierarchical selection (HS) models of the butterfly schema. The dichotomy between explanations of schemata, as rational generative structures or empirical associative-statistical phenomena, is explored. The tendency for congruence is supported by rationalist theories of psychology and cognitive science, such as Gestalt psychology and auditory scene analysis, examined in Section 3.2.1. Gestalt psychology and auditory stream segmentation are proposed to be limited conceptions when used as definitive laws of perception, as they are broadly incorporated in generative theories, such as *GTTM*, although can be incorporated as broad constraints of cognition.

The conditioning properties of the environment and culture, expounded in behavioural psychology and memetics, provide a basis for the top-down HS model of the butterfly schema. Attempts to explain musical structure through rational cognitive psychology or empirical associative methodologies alone, without integrating these perspectives, are limited. It is proposed that a synthesis of rational-universal-cognitive approaches and empirical-behavioural-memetic-cultural theories provide a richer understanding of musical systems. Behaviourism and memetics explain how grammars and schemata are contingent upon the cultural selection of their features, which follow a particular history of selection (examined in Sections 3.2.3 and 3.2.4). Behaviourism and memetics support the top-down HS model of the butterfly schema through their explanation of the mechanism of selection. The concept of style hierarchies is introduced in Section 3.2.5, showing how the features of grammars are related within a hierarchical system, and how universal and particular influences are integrated.

3.2.1 Gestalt Psychology and Auditory Scene Analysis

This subsection reviews Gestalt principles and auditory scene analysis, which have been influential in many spheres of music theory, music psychology, and computational musicology. Gestalt psychology broadly contends that wholes are perceived differently from the parts that comprise them (Wertheimer, 1923; Koffka, 1935). *Gestalten* are irreducible sensory phenomena. They are innate percepts that constrain the intake of information as top-down wholes, defining the mechanism of perception. Gestalt principles partly dictate the patterning in styles, since they limit how humans perceive and cognise musical structure (Meyer, 1956, 1989). While styles or grammars vary between cultures, the universal principles of Gestalt psychology, which constrain perception and cognition, are constant. The Gestalt principles of proximity and similarity are particularly important. Cooper and Meyer (1960) incorporate Gestalt principles to explain how grouping structure is organised:

> Grouping on all architectonic levels is a product of similarity and difference, proximity and separation of the sounds perceived by the senses and organized by the mind. If tones are in no way differentiated from one another – whether in pitch, range, instrumentation and timbre, duration, intensity, texture, or dynamics – then there can be no rhythm or grouping. (Cooper and Meyer, 1960, p. 9)

Deliège (1987) presents an empirical psychological study that supports the primacy of Gestalt principles as a basis for grouping. Deliège (1987) demonstrates that trained and untrained listeners can identify grouping boundaries according to the grouping preference rules of *GTTM*, which are based on Gestalt principles. Auditory scene analysis (Bregman, 1990) provides further support that musical structure is constrained by Gestalt principles. Bregman explores how the simultaneous separation of groups in texture, termed auditory stream segregation, is achieved. Perception segments textures into separate streams according to proximity, similarity, and good continuation. Auditory stream segregation is a universal constraint of psychology that governs musical structure, and so is an important concept for grammars (discussed in Section 3.1.4). Separate streams are perceived when notes of a single melody are separated enough though pitch distance to enable the fissure of the melody. Auditory stream segregation has been influential in theories of contrapuntal grouping, such as the polyphonic components of Lerdahl (2001) and Temperley (2001), and is presumably an implicit consideration in the perception of voice-leading schemata of Gjerdingen (1988, 2007). The schemata of Gjerdingen (1988, 2007) have multi-voiced textures, and so the identification of schemata hinges on the notion that individual streams are perceivable in texture (Gjerdingen (1994) examines the auditory perception of grouping). In short, Gestalt psychology and auditory scene analysis is the basis upon which many theories about the perception of musical structure rest.

A general criticism of the application of Gestalt principles to models of musical structure is that they do not actually explain the organisation of perceptual phenomena, but merely describe potential percepts. It is therefore possible that these perceptual phenomena might vary in different contexts (as suggested in Peel and Slawson (1984), Bod (2001b, 2001c, 2002), and Bruce et al. (2003, p. 134)). The application of Gestalt principles as broad constraints on musical structure is defensible because they have been demonstrated to be important constraints on perception (Tenney and Polansky, 1980; Temperley, 2001). However, when they are considered fundamental perceptual laws for parsing musical structure, as they have presumed to be in the field of vision (e.g., Koffka (1935)), and in generative theories of music (e.g., Lerdahl and Jackendoff (1983)), then this is more problematic. The principles of Gestalt psychology apply most comfortably to the natural categories of visual perception, but less so to the categories of musical structure (Bod, 2002), which represent artificial (not natural) categories. Even in the fields of visual and auditory perception there is sustained criticism about the explanatory power of Gestalt principles (e.g., van Noorden (1975), Carlyon et al. (2001), and Bruce et al. (2003)). An attempt to understand the mechanism behind Gestalt principles is sought in more recent ecological and physiological explanations of visual categories (Bruce et al., 2003), but such critiques are not common in perceptual studies of music.

Groups are often organised according to Gestalt perceptual principles, but this does not invalidate other possible cognitive organisation. That is, grouping structure might also be segregated through cognitive reflection on musical structure. Moreover, grouping structure can be influenced by cultural constraints. While the application of Gestalt psychology to music can be revelatory, the variety of grouping structures (in terms of length, internal leaps, accentuation, etc.) found in cultures is broad enough to suggest that Gestalt principles are not universal laws, but universal constraints. It is thus limiting to theorise about structure with the presumption that grouping structures are necessary perceptual *Gestalten*, since structure is the *product* of rational cognition and variable culture.

Recent investigations in the field of auditory perception provide further arguments against applying Gestalt principles too strictly (van Noorden, 1975; Carlyon et al., 2001; Fishman and Steinschneider, 2010). Temporal coherence refers to the notion that a sequence of pitches forms the impression of a linear connected series, as a function of pitch proximity and tempo. This is the reverse process to the fission of melodic streams that occurs in auditory stream segregation. The tempo of a sequence has to be reduced if there is a large separation in the frequencies between the pitches of a stream for it to be perceived as temporally coherent (Deutsch, 1982, pp. 119–120). At faster tempi, the sequence is fissured in cognition to form segregated streams. This means that Gestalt principles are anything but absolute laws of perception, since the outcomes of structure depend on the interaction between perceptual principles and other parameters, such as pitch proximity and tempo in this case.

Further to this argument, van Noorden (1975) has shown that between the threshold frequencies that listeners hear streams as a function of tempo, and the boundary frequency that listeners hear streams (at a range of tempi), there is subjectivity in the perception of streams, depending on the attention of the listener. Carlyon et al. (2001) likewise finds that learning is required to segregate streams, since the groupings determined by listeners changed over repeated listening — which also contradicts the emphasis on perception in Bregman (1990). Broadly, streaming requires the interaction between the top-down processes of cognition and the bottom-up processes of perception (Fishman and Steinschneider, 2010, pp. 238–239). As a universal perceptual constraint on musical structure, Gestalt principles can be incorporated into the theory of grammars, but only through considering its interaction with attention (cognition) and culture. The analysis of grouping is complex and resists the simple application of broad universal perceptual laws, as presented in many generative theories.

A tempered use of Gestalt principles is assumed in the tendency for congruence and model of the butterfly schema. In the butterfly schema model, Gestalt principles influence textural and metrical structure, but a flexible synergy of cognitive and cultural constraints is shown to determine musical structure. Challenging the credo that *Gestalten* are immutable laws does not deny their importance as organising principles of musical structure, but permits a greater variety of structures that can form in grammars.

3.2.2 Schema Theory and Prototype Theory

This section examines the theoretical foundations of schema theory and prototype theory. The crux of Gjerdingen's (1988, 2007) research programme rests on the advancements in schema theory in the work of Gibson (1969), Schank and Abelson (1977), Rummelhart (1980), and Mandler (1984), among others. Of the schema categories formed in cognition, some have definite boundaries but others are fuzzy (Rosch, 1978, p. 10; Zbikowski, 2002, pp. 36–40). Within categories (such as grammars and schemata) there are prototype effects. That is, there are asymmetrical distributions of members of a category (Lakoff, 1987, p. 41). Prototype effects observed in the domains of vision and language are primarily concerned with the perception of basic-level categories. While this is perhaps a different venture to the application of schema theory and prototype theory to the more complex artificial categories of musical structure, many of the observations and analysis might still apply. However, the theory of categorisation and prototypes of schemata in Rosch (1975a, 1975b, 1978) and Rosch and Mervis (1975) have been influential in cognitive science but seem not to have been absorbed into the prototype theory of Gjerdingen (1988). Gjerdingen (1988, pp. 94–121) argues that musical schemata are formed in cognition; they are categories with fuzzy boundaries that form around a central prototype. Gjerdingen (1988, p. 94) suggests that prototypes are the most ideal or typical type within a category, but arguably does not fully consider the complex nature of prototypes, which is now explored.

In contrast to Gjerdingen's interpretation of schema theory, the criteria for being a member of a category or schema does not directly correspond with those of being a prototypical member of that category or schema (Rosch, 1975a, 1975b, Rosch, 1978; Rosch and Mervis, 1975). A common, but seemingly false interpretation of Rosch's early work from the late 1960s and early 1970s (e.g., Rosch (1973)) is that the structure of prototypes directly mirrors the structure of categories (Lakoff, 1987, p. 42). Rosch's work of the late 1970s (e.g., Rosch (1978)) stresses that prototypes, which are asymmetrical patterns in schema categories, do not directly reflect category structure because additional factors constrain prototypes (also discussed in Lakoff (1987, p. 43)). Rosch et al. (1976, p. 382) argue that humans abstract categories from the environment in a way that is 'not arbitrary but highly determined'. '[A] category must have additional internal structure of some sort that produces these goodness-of-example ratings' (Lakoff, 1987, p. 45). That is, prototypes might not provide a characterisation of the internal structure of categories, nor do they mirror representations of categories (Lakoff, 1987, p. 43), although they might broadly correspond with categories. By contrast, Gjerdingen (1988) generally assumes a direct relationship between the structure of prototypes and the structure of categories. A prototypical 1–7…4–3 schema occurs when the arrangement of features are ideal according to the definition of the schema category.

Rosch (1978, p. 18) argues that the prototypicality of a schema depends on context and function. This means that the prototype is often more complex than the category in which it is found. Prototypes are asymmetrical sub-categories of schematic categories, which depend on those categories, but often concern deeper levels of analysis. Also, cognition might interpret categories in ways that are appropriate to the specific domain, being specific for vision, language, and music (Lakoff, 1987).

Cognitive scientists have examined the causes of prototype effects in various domains (summarised by Lakoff (1987, pp. 12–46)). Berlin and Kay (1969) found that far from being arbitrary representations, basic colour categories identified by humans relate to real categories in the world. However, while the focal colours of these categories are universally agreed upon (white, black, red, green, yellow, blue, brown, purple, pink, orange, and grey), the boundaries between colour categories, where non-focal colours reside (such as scarlet or saffron), are fuzzy. Prototypical focal colours and basic colour categories are defined through the interaction between biology, embodied cognition, language systems, and the cultural environment. The experience of focal colours is embodied in cognition, through the evolved eye cells in humans and primates. These focal colour cells are the universal constraints of neurophysiology that define colour categories. Kay and McDaniel (1978) examine the vision cells of monkeys to explore how colour is perceived. The base firing rate of vision cells in monkeys are similar to humans, and are set to the primary colours, blue, yellow, red, and green. The perception of colour changes through alterations of the firing rate of the cells. This means, broadly speaking, that cell firing rates that are primed for primary colours also constrain the perception of non-primary colours. Thus, focal colours and nonprimary colours are partly determined by embodied cognition. The boundaries of a colour category are fuzzy because of the lack of agreement between people owing to the influence of language and culture, even though the categories are broadly universally constrained.

Cultures disagree on whether colours belong to one category or another because they are governed by embodied cognition, language, and culture. 'Color categories result form the world plus human biology plus a cognitive mechanism that has some of the characteristics of fuzzy set theory plus a culture-specific choice of which basic color categories there are' (Lakoff, 1987, p. 29). The analysis of Kay and McDaniel (1978) shows that colour categories are dependent on embodied cognition. They are universally constrained in much the same way as pitch-space constraints and Gestalt constraints govern the formation of universally generalisable rhythmic and pitch structure of generative theories. These investigations of visual colour categories show that categories are a complex admixture of universal and particular constraints, much as categories are described in generative theories of music, such as *GTTM*. However, in artificial categories, such as in music, there is more reason to suggest that universal constraints and particular constraints combine in complex ways, although a deep understanding of this complexity is probably beyond the scope of current science.

Gjerdingen (1988, p. 94) posits that prototypes are simply the most typical or ideal examples of a category, similar to explanations in the early research of Rosch (1973). A discussed, Lakoff (1987, p. 43) points out that assuming that the structure of prototypes is indicative of categories is a common misapprehension. Indeed, the examination of prototypes in Rosch (1978) and Lakoff (1987) accords with the present theory of the tendency for congruence, because it shows that categories and prototypes are formed in combination with cognitive (embodied) universal constraints. In a different vein, Lakoff (1987) also presents a generative theory of semantics, positing that language structure is constrained by embodied cognition and culture. Schematic categories are similarly formed in this theory through the interaction of cognition and culture (Rosch, 1978; Lakoff, 1987). Gjerdingen (1988) does not consider embodied cognition, and assumes that the internal structure of categories (prototypes) mirrors the overall structure of the category, since schemata are the most ideal or typical examples of a category in Gjerdingen (1988, p. 94), which is an overly simplistic application of the schema principle.

The distinction between the term 'grammar', as a term for global structure in music (used in generative theories), and 'schema', as a term for local structuring (in associative-statistical theories) is not held in the psychological literature. Indeed, the two terms are used interchangeably for local and global, abstract and concrete, or universal and particular patterns. Also, grammars and schemata are often defined as mental theories or abstractions that are used as procedural tools for interacting with the environment. Rumelhart (1975, p. 213) describes a story grammar that follows a stereotypical patterning in the ordering of its features. Each segment of the story has a function and purpose relative to its context in the unfolding structure (Rumelhart, 1975, p. 217). This corresponds with the conception of the Markovian schema thread ('*il Filo*') in Gjerdingen (2007, pp. 369–397). Indeed, schemata and grammars can both follow a general or abstract order and pattern, and can be used in a variety of contexts, local and global.

Rumelhart (1980, pp. 33–40) argues that schemata are functional and that they are broad models of the environment that often sacrifice detail for generality, permitting greater flexibility of abstraction. That is, in this definition, schemata, like grammars, permit combinatorial freedom. This is similar to the definition of grammars presumed in generative theories, but contrasts with the specificity of voice-leading schemata in Gjerdingen (1988). Schemata are knowledge of environmental patterning that provides a map by which humans can understand and interact with the environment. They are not just interpretations of the world, but permit solving problems in it (Rumelhart, 1980, p. 55). Accordingly, they are ways of theorising about the world that reveal how the mind interprets the environment. Rumelhart (1980, pp. 40–41) lists the major characteristics of schemata, reinforcing the notion that schemata are more than just abstractions of the environment, but are functional or operational theories:

1. Schemata have variables.
2. Schemata can embed, one within another.
3. Schemata represent knowledge at all levels of abstraction.
4. Schemata represent knowledge rather than definitions.
5. Schemata are active processes.
6. Schemata are recognition devices whose processing is aimed at the evaluation of their goodness of fit to the data being processed.

(Rumelhart, 1980, pp. 40–41)

Gjerdingen (1988) focuses on the associative relationships between schema features, suggesting that schemata are mainly abstracted patterns from the environment. However, there is little consideration about how they are active processes used by cognition to interact with the environment. While the perception of the schema life cycle (as described in Gjerdingen (1988, pp. 99–106)) is operational in a sense, this application of schema theory does not fully explain why schemata form, whether through culture, universal perceptual principles, or real-time listening. The retrodictive analysis of the life cycle of the 1–7…4–3 schema describes its historical evolution, rather than showing how it is cognised *prima facie*. Schemata are not fully examined in terms of their implicit cognitive structure or in terms of the cultural constraints acting on them. Therefore it is unclear how humans are supposed to understand schemata on first exposure, as an eighteenth-century listener would, or a listener from any historical period. While universal embodied cognition influences the formation of categories and prototypes (as explained in Rosch (1978) and Lakoff (1987)), Gjerdingen (1988) simply defines schemata as fuzzy cultural patterns that are present in structure and which evolve through time.

Gjerdingen (1988, p. 99, 2007, p. 11) broadly contends that schemata mirror the empirical distribution of musical structure in the real world, which raises questions about why a schema theory framework is required for explicating the structures concerned. That is, if schemata are governed by statistical inferences of culture, a theory centred on cognition is perhaps not necessary. To compare Gjerdingen (1988) with the schema theory of Piaget (1952), for example, Piaget's schema theory explains developmental schemata as mental strategies. As children develop, various schemata are invoked that help them to understand the world, appropriate to their stage of development. In this way, Piaget shows that schemata help the child to understand the world at various stages of development. In doing so, it is more convincingly a theory of cognition. These schemata, in a sense similar to the frames of mind in Minsky (1986), explain the development of the child as it becomes increasingly sophisticated. In Gjerdingen (1988), however, it is arguably not necessary to invoke the schema concept, because this portrayal is primarily concerned with categories of culture, not cognition.

3.2.3 Selection by Consequences

This subsection examines how theories of selection explain how culture constrains the formation of the butterfly schema and the Classical grammar. Behaviourism and memetics concern the selective mechanism of culture. Both give primacy to the role of the environment and culture in shaping human behaviour. Behaviourism subscribes to the universal Darwinian notion that selection extends to domains other than biology. However, selection in behaviourism generally occurs at the level of human behaviour, whereas it occurs at the level of culture in memetics. This subsection and the following subsection (Section 3.2.4) provide overviews of behaviourism and memetics, respectively. Selectionist theories of culture are a framework for the top-down HS model of the butterfly schema and Classical grammar (presented in Section 5.3), showing how features of grammar exist in a particular order of dependency.

The mechanism of cultural selection is a contrary explanation of behaviour and culture than presented in many rationalist accounts. Prior to the cognitive revolution of the 1950s and 1960s, which was partly initiated by the generative enterprise in linguistics (e.g., Chomsky (1957, 1959)), behaviourism occupied a central place in psychological discourse. Behaviourism fell sharply from favour in many disciplines mainly after developments in computer science, generative linguistics, and cognitive science. However, it provides an important theoretical opposition to the rationalist philosophy of mind, explaining the empiricist forces of environment and culture on human behaviour. In behaviourism, the behaviour of an organism is explained in terms of the conditioning stimuli in the environment. The primary mechanisms of behavioural analysis are *classical conditioning*, where beings give conditioned responses to antecedent environmental stimuli (Pavlov, 1927), and *operant conditioning*, where behaviour is shaped and maintained by its consequences

(Skinner, 1953, 1957, 1981; Baum, 1994). The latter process is of most interest to the present model of the butterfly schema because it shows the primacy of past consequences for selecting subsequent behaviour and culture.

The notion of selection is most comprehensively explained in terms of the 'selection by consequences' of Skinner (1981). This broadly universal Darwinian perspective unifies biology, behaviour, and culture through the mechanism of selection:

> Only past consequences figure in selection. (i) A particular species does not have eyes in order that its members may see better; it has them because certain members, undergoing variation, were able to see better and hence were more likely to transmit the variation. (ii) The consequences of operant behaviour are not what the behaviour is now for; they are merely similar to the consequences which have shaped and maintained it. (iii) People do not observe particular practices in order that the group will be more likely to survive; they observe them because groups which induced their members to do so survived and transmitted them. (Skinner, 1981, p. 503)

At the outset, selection by consequences is confusing because behaviour and culture are generally thought to have rational antecedents. However, a consideration of the selection history of human behaviour suggests that although humans appear to be initiating agents of behaviour, their intentions are often inconsequential in terms of the net outcomes of behaviour. This is similar to biological evolution, where a species does not adapt to an environment, but the consequences in the environment select adaptive traits. So too with human behaviour and culture, a person or culture does not adjust to a situation but consequences in the environment select and shape the behaviour and culture. The behavioural conception argues that cognition merely appears to be the ultimate cause of behaviour, but the environment or culture is the ultimate cause.

Selection in the domains of biology, behaviour, and culture work together hierarchically and redundantly (Skinner, 1981, p. 501). All types of selection are ultimately reducible to natural selection. Baum (1994, p. 78) explains this interaction, positing that '[j]ust as differences in reproductive success (fitness) shape the composition of a population of genotypes, so reinforcement and punishment shape the composition of an individual's behaviour …. Where inherited behaviour leaves off, the inherited modifiability of the process of conditioning takes over'.

A selectionist account of culture would mean that unique human ideas do not influence culture, but rather the consequences in culture over time shape those ideas (Skinner, 1981, pp. 501–504). Behaviourism presents a mild contrast with memetics, because the latter posits that cultural material can override operant conditioning and natural selection. The memetics viewpoint is plausible, but since culture operates at a more particular level of selection to operant behaviour and evolution, the frequency of memetic selection overriding natural selection and behavioural conditioning is lower. (Blackmore (1999) convincingly argues the case for situations where cultural selection dominates natural selection and operant conditioning.)

Understanding language as operant behaviour ('verbal behavior' in Skinner (1957)) or cultural selection diverges strongly from viewing it as a cognitive capacity (Chomsky, 1959, 1965, 2006). Chomsky (1957, p. 17, 1959, 2006, p. 98–99, p. 177) argues that language (and perhaps, by extension, music) is not primarily designed for communication, but for internal thought (an I-language, not an external E-language). The Aristotelian notion that language is 'communication in sound' is inaccurate in this view, since communication is a symptom of the underlying generative capacity for language. Language is a uniquely human recursive system (Humboldt, 1836/1999), capable of discrete infinity (as shown in the 'universal grammar' of Chomsky (1957)). Recursion is the mathematical procedure for embedding phrases within other phrases. Recursion and discrete infinity enable infinitely complex edifices of words and sentences within a hierarchical structure using the discrete building blocks of words. To illustrate, six-word and seven-word sentences are commonly generated in language, but six-and-a-half word sentences are not (Chomsky, 2000, p. 51). These attributes of language are proposed to be universal, although debate continues to run concerning the definitions of language, its purpose, and structure (as found in the diverging views of Chomsky et al. (2002) and Pinker and Jackendoff (2005)). Languages have been documented where there is no recursion, such as that of the indigenous Pirahã tribe of the Amazon (Everett, 2005). Such anomalies are rare and perhaps do not affect the potential generality of universal grammar, since exceptions do not prove that a generative cognitive capacity does not exist, but only that this capacity is not utilised in certain circumstances.

Cultural selection (and operant conditioning) contradicts the notion of a generative capacity of music because the cognitive agency of the composer is perceived as being insignificant compared to the selective contingencies of environment and culture. A behaviourist view of music might explain musical structure as the consequence of a complex system of operants, which are behavioural responses to environmental stimuli. Features of grammars are likewise selected in a chain, where each feature is a consequence of previous features. Thus, from a behavioural perspective, the cause of musical behaviour is primarily external to cognition. Notwithstanding, the behavioural approach is limited precisely because it does not offer an explanation of cognition, which must also constrain musical structure.

Setting aside this limitation, omitting a consideration of the cognitive influence on musical structure has a key advantage. A behavioural account of musical structure is tractable and verifiable, whereas mentalistic notions of musical structure are often unquantifiable. When music is viewed quantitatively, as a product of culture, elegant and testable scientific theories can be constructed. Meyer (1989, pp. 142–150) views culture as a main influence on musical structure, positing that music is mainly contingent upon environmental and cultural rules. Meyer (1989, p. 142) does not espouse a 'radical' behaviourist view, because he admits that human cognition is also significant, but contends that culture is a primary constraint on composers, and that there are very few conscious 'choices' in the composition of musical structure. Therefore Meyer is a behaviourist in a sense, since operant conditioning and cultural influences are presumed to be the key shapers of musical structure. Meyer (1956, 1973, 1989) also includes rationalist explanations of musical structure, such as Gestalt principles, but generally does not give attention to generative theories.

While cultural selection constrains musical structure in this thesis, the tendency for congruence — an indirect product of cognition — forms the implicit structure of music. Therefore, a tempered form of cultural selection is assumed in the top-down HS model (in Section 5.3), explaining the formation of, and dependency between, musical features in the Classical grammar and butterfly schema, because it is accepted that cognition also has an important role. There is no attempt to explain the entirety of behaviour as a product of operant conditioning and cultural selection, as is attempted in behavioural psychology (Skinner, 1953, 1957, 1981). A selectionist account of musical structure is necessary for understanding the cultural conditioning of features.

3.2.4 Cultural Transmission

Of the theories of cultural evolution and selection, memetics has perhaps been most fully elaborated in music theory (see Jan (2007), (2013)). Similar to behaviourism, memetics counters the notion that humans are intentional beings free to design new, or novel artistic creations. Memetics considers culture to be a more influential constraint on musical structure than cognition, limiting or denying the agency of humans as creators (as argued in Dawkins (1976), Blackmore (1999), and Jan (2007)). Memetics is an approach to understanding culture that gives primacy to the evolutionary algorithm of variation, selection, and heredity (replication). These aspects of memetics are incorporated into the top-down HS model of the butterfly schema.

Memes are units of culture of arbitrary length, acting in a similar way to genes, replicating themselves in culture and brains (Blackmore, 1999, pp. 7–8). Memetics, in its broadest conception, is concerned with the replication or imitation of cultural units (Dawkins, 1976, p. 192). Memes operate in a variety of substrates, they have no purpose, agency, or goal, and are replicators which are selected against each other. Memes are selected for their longevity, fecundity, and copying-fidelity (Dawkins, 1976, p. 194). Memetics views culture through the meme's-eye-view (Blackmore, 1999). From this perspective, '[m]emes have no foresight, they do not look ahead or have plans or schemes in mind' (Blackmore, 1999, p. 13). Accordingly, memetics draws similar conclusions to the framework of selection by consequences (Skinner, 1981). Both behaviourism and memetics generally contend that environment and culture are more significant influences on human behaviour than mentality or cognition.

Blackmore (1999) theorises that humans have evolved in order to spread memes, as custom-designed imitation vehicles.[21] The evolution of human cognition is thus proposed to be the product of cultural selection. Imitation is a 'good trick' for humans because it is economical in time and less dangerous than procuring techniques and knowledge individually. This capacity for imitation improves fitness and creates a selection pressure to imitate, which bootstraps evolutionary development (Blackmore, 1999, p. 75). Blackmore (1999, p. 33) argues that '[t]he human language capacity has been meme-driven, and the function of language (if it has any) is to spread memes'. However, the suggestion that language might have developed as a result of the selection pressure to imitate is contentious. The reverse of this hypothesis is also feasible: the increasing cultural imitation through history is the consequence of evolving cognition. Moreover, the cause of increasing cultural activity through the Upper Palaeolithic period (50,000–10,000 BP (before present)) might be due to the shift of general human sophistication during this time, and, more controversially, the seemingly sudden emergence of the generative capacity in humans (as claimed by Chomsky (2006, pp. 173–185)). Indeed, although the memetic perspective is an attractive one, the generative capacity for language in humans is not easily explained as an evolutionary adaption. Chomsky (1966, pp. 173–185) contends that the generative language capacity is unlikely to have been naturally selected because there are no legitimate selection pressures that explain mathematical recursion.

As discussed, memetics and behaviourism generally see human intelligence and rationality to be on the reins of culture. Both theories avoid mentalistic explanations of internal events, which are viewed as unscientific. Conjecture about inner selves invites further conjecture of inner selves, which results in an infinite regression (Pagel, 2012, p. 270). Memetics considers internal events only in the sense that memes are cultural units that are replicated in brains, therefore avoiding epistemological problems of defining knowledge and the *problem of induction*. Dennett (1978b, p. 58) points out that rationalism must necessarily be sidestepped in memetics, suggesting that, '[s]ince psychology's task is to account for the intelligence or

[21] Since the publication of Blackmore (1999) it is has become more widely accepted that Neanderthals are cognitively more similar to homo sapiens than was previously thought, owing to the knowledge that they have interbred, and share DNA (Elias, 2012). This partly contradicts Blackmore's (1999) claim of the necessary relationship between the development of human cognition and the memetic transmission of culture, since the assumption that Neanderthal culture was less sophisticated than early human culture, upon which her claim was partly based, must also be challenged.

rationality of men and animals, it cannot fulfil its task if anywhere along the line it presupposes intelligence or rationality'. Since ideas about mentality and consciousness struggle for scientific bearing, science must examine the external manifestations of cognition (Dennett, 1991, p. 41). While neither classical conditioning nor operant conditioning can account for imitation, memetics potentially provides a scientific framework (Blackmore, 1999, p. 45) because it looks at cultural units directly. However, it is not an exact science because the cultural topography is complex and less particulate than its primary model, genetics. While behaviourism concerns human behaviour in short time frames, memetics generally concerns cultural transmission that takes place over longer time frames. Even though memes seem to act independently and selfishly, they are dependent on other memes for their replicative success. Co-indexation is the combining of memes with similar bases. The survival of one meme is dependent on other memes with which it is co-indexed (Jan, 2013, 2007; Blackmore, 1999; Dawkins, 1976). This involves the selection of memes by other memes in memeplexes (Dawkins, 1999, pp. xiv–xv).

This concept of cultural selection in memetics contradicts the hegemony of natural selection over operant conditioning and cultural selection suggested in Skinner (1981) (where culture works redundantly with the more powerful forces of operant conditioning and natural selection). In Skinner (1981), culture is selected only if it permits the survival of the individual. However, in memetics, culture can counteract the conditioned behaviour of the individual (which is a central claim of Blackmore, 1999). While Skinner (1981) does not argue that culture is invariably contingent on operant conditioning (since he is undoubtedly aware that there are instances when culture is detrimental to the survival of the individual), Skinner proposes that when culture is viewed sufficiently long-term, it is generally more often contingent on the outcomes of natural selection and operant conditioning.

Memetics is thought to be either analogous with biological evolution (Dawkins, 1976; Blackmore, 1999), or parallel to it (following the hypothesis of universal Darwinism (Jan, 2013, p. 12, 2007, p. 15)). Dawkins (1976, p. 189), who conceived the concept of the meme, cautions that memetics is a theory that is merely analogous with genetics. Blackmore (1999, p. 66) claims that beyond their replicative properties the analogy between genes and memes is indefinite. The stronger hypothesis, that memetics is parallel with biological evolution (considered specifically in music in Jan (2013)), argues that various human structures are primarily caused by memetic replication, rather than cognition. This stronger hypothesis diminishes the rational agency of cognition, which is controversial because it lowers the value of humanly created structures, such as music. While memetics offers a more direct framework to deal with the action of culture than does behaviourism, the lack of attention to cognitive causes in both approaches presents a major limitation to understanding musical structure.

Dennett (1995, pp. 352–360) explores the concept of selection in behaviour and culture, following a similar line of argument to the notion of a 'science of human behaviour' in behaviourism (cf. Skinner (1953) pp. 11–22).[22] Dennett argues that memetics might be able to explain human cognition. Dennett (1991, pp. 182–226) examines various internal selection mechanisms, showing that selection and replication occur at a number of distinct levels in cognition. Dennett shows that these types of internal selection form a 'Tower of Generate-and-Test' that are graded in terms of their usefulness for organisms to survive in the environment (Dennett, 1995, pp. 373–381).

Natural selection in biology seems to be the most fundamental mechanism of selection, but offers no ability for a postnatal design-fixing, required for organisms to coherently correspond with environmental conditions (Dennett, 1995, pp. 374–375). Dennett (1995, p. 375), after Popper and Eccles (1977), describes humans (and animals, but to a much lesser extent) as 'Popperian creatures'. They have the ability to test hypotheses internally before committing to action. This mechanism of hypothesis-testing could work contradictory to natural and cultural selection because potential contingencies that are selected internally could counter the contingencies offered by the environment and culture. Cognitive scientists and psychologists might disagree in principle with this notion of internal selection since cognition could be argued to be more refined than a brute selection mechanism (cf. Pinker (1997, 2002)).

[22] A scientific approach to human behaviour and culture is also found in Marxist materialism, the cultural materialism of Harris (1979), and many other empiricist philosophies and studies, such as Nietsche's *Beyond Good and Evil* (1886/2002) and the writings of Locke and Hume.

Dennett (1995, pp. 377–378) illustrates an internal selection mechanism that is a unique capacity of humans, and which supports the primacy of cultural selection. Following the theory of culture in Gregory (1981), Dennett posits that humans are 'Gregorian creatures', because they are able to import, mainly through language, 'information tools' of culture and technology that can be applied to environment problems. This lends broad support for cultural selection because it demonstrates that complex bodies of information are exchanged through brains and culture that are not necessarily rationally interpreted, even though they seem to empower individuals intellectually. However, in mild contrast to memetics, where culture can dominate natural selection, Gregorian creatures are generally at the behest of natural selection (although not necessarily so). That is, humans are naturally selected if they import useful information tools, and not naturally selected if they do not.

Behaviourism and memetics are thus useful in this dissertation for providing an account of the mechanism of selection. Memetics also considers the notions of variation and replication, which are assumed in the top-down HS model of the butterfly schema. While these processes can describe the external cultural influences on musical structure, they cannot account for the cognitive influences, which are definable through the tendency for congruence. The use of the selection mechanisms in behaviourism and memetics is incorporated into the top-down HS model. Although the selective mechanism of culture acts on musical structure, the influence of the tendency for congruence is significant also. Thus a combination of empirical and rational theories and methodologies is shown to be necessary for explaining musical structure.

3.2.5 Style Hierarchies in Music

This section provides an overview of musical theories and models of hierarchical systems in music. Meyer (1989), Nattiez (1990), and Jan (2007) present general style hierarchies that show the interrelationships between universal laws and particular rules. Meyer (1989, pp. 13–24) presents a general qualitative model of style hierarchy, and Nattiez (1990, p. 136) shows a general style hierarchy in graphic form. The aim of this section is to provide an overview of hierarchical systems, specifically in music theory. (Systems theory (von Bertalanffy, 1968), selection hierarchies, used in multi-level selection theory (Okasha, 2006, 2012), and causal systems (Ellis, 2005, 2008), are examined in Section 5.3.)

Meyer (1989) introduces three main types of strata in hierarchies: universal laws, rules, and strategies. Laws are universal to all musical cultures, rules are intracultural (particular), and strategies are procedures for determining how rules are instantiated. Meyer (1989, p. 13) posits that laws, in the first stratum, are of three types: physical, physiological, and psychological. Similar to the views of Meyer (1956, 1989), Lerdahl and Jackendoff (1983), Temperley (2001), Lerdahl (2001), and Lerdahl (2011), the most significant of these are psychological laws, such as those of Gestalt psychology (Wertheimer, 1923; Koffka, 1935). Treating Gestalt principles as invariant laws is problematic because universal essences that are consistently applicable across cultures probably do not exist (as discussed in Section 3.2.1). As Lerdahl (2001, p. 381, 2012) argues, cultures have the potential to create exceptions to universals. Therefore incorporating Gestalt principles as psychological laws places unnecessary strictures on the possible forms of structure that can form in grammars. Rather, universals must merely be constraints on musical grammars.

The intracultural rules are the second type of stratum in Meyer's style hierarchies. Of the intracultural rules, Meyer (1989, p. 17) distinguishes three types: dependency rules, contextual rules, and syntactic rules. Dependency rules are contingent upon the syntactic rules of another parameter (Meyer, 1989, p. 18), such as in medieval organum, where the rules of harmony are governed by the rules of voice leading. However, in other genres and grammars this order can be different. Contextual rules are those that are ascribed to particular points of pieces, such as cadences. They involve rules that are distinct from the normative global syntax (Meyer, 1989, p. 18), whereas syntactic (grammatical) rules are the main rules of the particular system (Meyer, 1989, p. 19).

Strategies are situated in the third type of stratum, which involve the orientation of the intracultural rules in cultures. 'For any specific style there is a finite number of rules, but there is an indefinite number of possible strategies for realising or instantiating such rules' (Meyer, 1989, p. 20). Strategies are similar to dependency rules and contextural rules but concern the more general aspects of structure, involving the manipulation of pre-existing rules, but not changing the rules themselves. Dependency rules, contextual rules, and strategies are types of meta-rules that orient intercultural rules, providing information about how those rules are instantiated. Meyer's (1989) conception of style hierarchy therefore not only involves the interaction of syntactic rules but also incorporates information about the implementation of rules. This is because syntactic rules occupy not only a particular place in a certain hierarchy in times and places, but that the same rules are used in various other times, places, and contexts. How these rules are implemented depends on the grammar and context.

This qualitative explanation of style hierarchy in Meyer (1989) is a broad-stroke integration of universals and particular categories, including aspects such as intraopus rules (pertaining to a particular composition), composer rules, dependency rules, contextual rules, and strategies. In the top-down hierarchical model of the butterfly schema, many of these aspects are not included because they are deemed not to be quantifiable. Moreover, it is not necessary to incorporate meta-rules that explain how features interact because the mechanism of hierarchical selection is sufficient to show causal relationships between features. The model of the butterfly schema and the Classical grammar, presented in Section 5.3, is also different to the generalised style hierarchies of Meyer (1989), Nattiez (1990), and Jan (2007) because a particular cultural grammar and schema is under analysis.

Notwithstanding the above omissions, the style hierarchy explored in Meyer (1989) and Nattiez (1990), and the memetic hierarchy in Jan (2007), provide a foundation for a theory of cultural selection and evolution. Jan (2007) conceives memetic hierarchies (similarly to Meyer and Nattiez), as systems that are governed by universals, with each level constraining the production of culture on the next layer. In Jan (2007), memes are represented at the bottom of the hierarchy, which are constrained by laws, rules, dialect, idiom, and intraopus styles. Although selection between levels of the hierarchy is not directly proposed, it might be assumed. In Figure 3.1, Jan (2007) compares biological and memetic hierarchies to show the parallels between the systems. This illustrates that the processes operating in biology, namely variation, selection, and heredity, operate in music also.

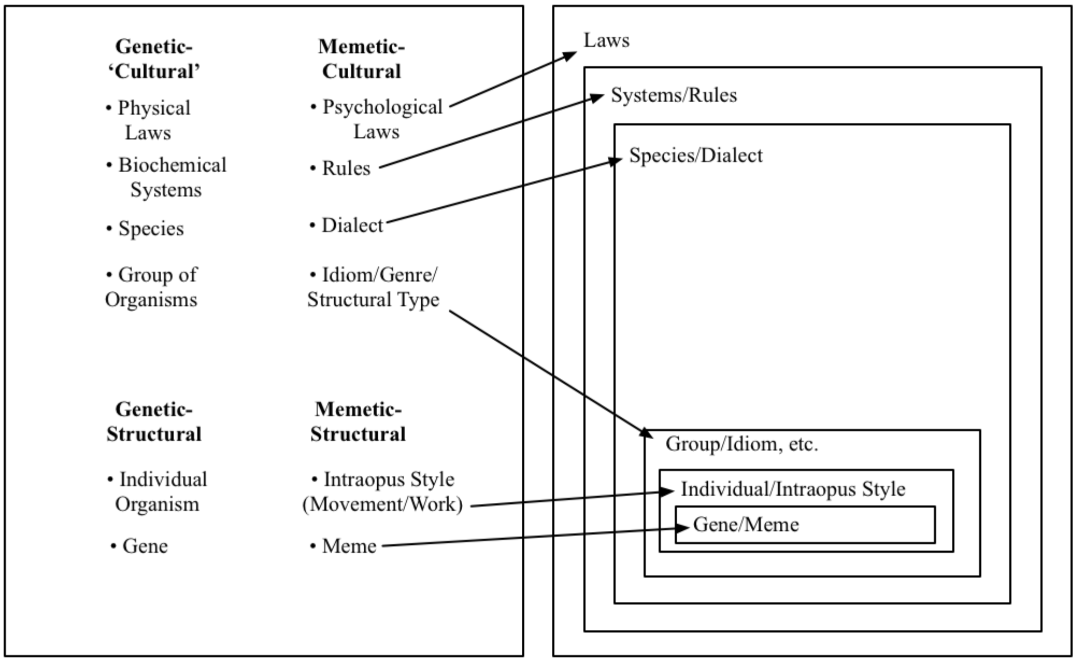


Figure 3.1 Correspondences between biological and memetic hierarchies (Jan, 2007, p. 106).

An important parallel between the two hierarchies in Figure 3.1 are the boundaries between physical and psychological laws and biochemical and cultural rules. The hierarchy tacitly diminishes the view of humans as intentional or rational beings, because universal laws (or constraints) and culture largely constrain the design of musical systems. The higher levels in Jan's (2007) hierarchy represent memes or group of memes that are replicated within the stratum.

This approach to hierarchy shows that features in the levels of hierarchical models can constraint other features in other levels. This type of cultural causation is proposed to differ to how listeners construct these features in real-time, which is generally characterised by bottom-up causation (shown in Section 4.1.2). The top-down HS model of the butterfly schema is based on multi-level selection theory (Okasha, 2006, 2012) and hierarchical causation (Ellis, 2005, 2008). Top-down HS generally works similarly to how these aspects are presented in Meyer (1989), Nattiez (1990), and Jan (2007), where larger wholes constrain the formation of constituent parts.

Hierarchical selection (HS) works in a similar way to allelic exchange in memetics, where memes are selected for in particular loci by other memes (Jan, 2007, 2013). The concept of co-indexation and juxtaposition (or 'crossing-over) of memes in memeplexes (Jan, 2007, 2013) analogises with hierarchical selection because memes (features) are constrained, by other memes at higher levels of the hierarchy. In the same way, each meme of the Classical grammar and butterfly schema is constrained by features at higher levels. By contrast with the general hierarchies given in Meyer (1989), Nattiez (1990), and Jan (2007), the top-down HS model of the butterfly schema will show a detailed cross-section of a discrete section of a particular grammar — instrumental music of the Classical period. This is a specific and novel conception of hierarchical conception of HS.

3.3 Conclusions

Computational, psychological, and cultural theories and models present various insights that inform the concept of the tendency for congruence and the multiparametrically congruent and top-down HS models of the butterfly schema. The implementations of harmonic, metrical, and grouping computational systems show the implicitly congruent (and culturally conditioned) formal structural relationships of musical structure. A limitation of these models is that they often do not fully consider multiparametric relationships, which could provide greater understanding of the cognitive and cultural influence on musical structure. The multiparametric approaches of Hamanaka et al. (2006) and Marsden (2005, 2007, 2013) are interesting because they epitomise the strengths and limitations of generative approaches. They characterise nuances of grammatical relationships, however, particular features, such as intraopus style and composer style, are generally too fine-grained to permit a formal analysis. It is also significant that they demonstrate that explanations of generative procedures in cognition are elusive.

A number of algorithmic models are interesting because they address the challenge of explaining the implicit tendency for congruence in musical structure. Fewer inductive or associative computational models are multiparametric because it eludes analysis. Generalisable frameworks that formalise the multiparametric cognition of musical structure are probably impossible to implement. Key contributions that do consider multiparametric relationships include Tenney and Polansky (1980), Marsden (2007, 2010, 2013) and Hamanaka et al. (2005, 2007). Tenney and Polansky show the relationship between the Gestalt principles of similarity and proximity. It can be extrapolated from their model that pitch similarity is closely connected to rhythmic proximity in cognition. Tenney and Polansky (1980) is an operational theory that foreshadows the preference rule system of *GTTM*.

The polyphonic grouping component of Temperley (2001) also considers relationships across parameters, to an extent, showing how rhythmic congruence produces metre. However, Temperley conceives polyphony in terms of homorhythmically unified groups, which is problematic because polyphony, partly by definition, can include non-isochronous textures. That is, the onsets of voices in polyphonic textures can be much more varied than homorhythmic textures. Notwithstanding, this approach to polyphony greatly influences the present work because in the Classical style contrapuntal textures are generally homorhythmic, and so textures and metrical structure can be grouped holistically, termed textural grouping (espoused in Section 5.1.2). The criticisms of Gestalt principles in Peel and Slawson (1984) and Bod (2001b, 2001c, 2002) are significant because they show the limitations of perceptual constraints for understanding musical structure. Bod (2001c, 2002) shows that Gestalt laws of proximity and similarity are not as significant as induced knowledge for the parsing of grouping structure.

This chapter raised a general criticism concerning inductive computational models. Harmonic, grouping, and metrical structure models often provide generalised and normative descriptions of musical structure, rather than offering procedural knowledge of the cognition of musical structure. Inductive frameworks cannot describe the process of cognition, but use pattern-matching techniques or key profiles to describe structure. These processes have been argued throughout this chapter to be limited mechanisms for explaining the cognition of musical structure. The recombination machines of Cope (1996; 2001) attempt to show how musical structure is cognitively organised, but rely on network relationships that do not fully explain the cognitive processes involved. Moreover, distinctions between composer style, intraopus style, and the idiom or grammar is vague, since inductive methods do not explain the interaction between the elements of networks (as argued in Cohn and Dempster (1992)). In Cope (1996, 2001), the relationship between signatures (bottom-up) and the hierarchical superstructure (top-down) is not clearly defined, and there may be no feasible way to show how this might be done.

Universal psychological and psychoacoustic constraints are integral for understanding musical structure, and are significant for demarcating the boundary between universal and particular structure in grammars and schemata. Psychological disciplines, and Gestalt psychology in particular, show the significance of universal constraints for modelling structural relationships. Gestalt psychology and schema theory are useful but limited ways of understanding musical cognition. While Gestalt psychology provides prospective strategies for modelling musical perception, it does not account for the variety of grouping structures in cultures. Gestalt principles are universal constraints that shape grammars and schemata but can be overridden by cultural constraints.

Schema theory and prototype theory, as often traditionally conceived in cognitive psychology (e.g., Rosch (1978)), differ in important ways from their use in Gjerdingen (1988). Whereas Rumelhart (1980) emphasises the procedural aspects of schemata, and Rosch (1978) and Lakoff (1987) show that prototypes are asymmetric with category structure, Gjerdingen (1988) construes schemata as salient and particular cultural patterns, without an explanation of the internal structure of schema categories. Also, Gjerdingen posits that prototypes are merely ideal members of schema categories, without an explanation of how categories are formed. Gjerdingen (1988, 2007) does not consider the effect of embodied cognition on musical structure that is important to prototype and category structure.

Behavioural and cultural theories, notably behaviourism and memetics, provide a framework for a theory of top-down HS. Behavioural and memetic theories comprise selective mechanisms that explain the influence of culture on musical structure. Style hierarchies show how style is generally structured into universal and particular strata. The style hierarchies of Jan (2007, 2013) have a degree of dependency between strata, but the top-down HS model of the butterfly schema will show that strata are highly interdependent. Memetic and behavioural theories of musical transmission must be counterbalanced with an understanding of universal physical, physiological, and psychological constraints, which are here proposed to be reducible to the tendency for congruence. Therefore the top-down HS model of the butterfly schema incorporates the tendency for congruence, which is a primary cognitive constraint on musical structure.

Chapter 4: Multiparametric Congruence in Butterfly Schemata

Chapter 1 explained how congruence might be manifest in particular grammars. Uniparametric congruence (UC) is the stability between elements of a single parametric feature; MC is the simultaneous occurrence of more than two uniparametrically congruent features; and the rule of multiparametric congruence (MC) is the tendency for congruence in grammars in particular contexts (discussed in Section 1.3.2). Chapters 2 and 3 argued that the rule of MC is tacit in proto-generative and generative music theories, and is a predominant consideration in computational models and psychological disciplines.[23] This chapter examines MC as the implicit organisational structure of the Classical grammar and butterfly schema, which is intuitively perceivable through the multiparametrically congruent interaction between abstract uniparametrically congruent features (outlined in Section 1.3.1). The butterfly schema is broader and more abstract than the concrete models of voice-leading schemata in Gjerdingen (1988, 2007), but is more particular than the generalised portrayal of tonal grammars in generative theories (e.g., Lerdahl and Jackendoff (1983), Temperley (2001)).

Section 4.1 examines how the highly abstract tonal and metrical structures of grammars can be diversely formed owing to the interaction between various abstract multiparametrically congruent features in particular grammars. A preliminary multiparametrically congruent model of the butterfly schema is presented in Section 4.2 to show how the particular abstract uniparametrically congruent features of the Classical grammar interact. (A preliminary model is required because the UC of each feature requires more detailed quantification, presented in Section 5.1). Section 4.2 explores the abstract uniparametrically congruent features of the

[23] The notion of congruent cognitive representations of musical structure perhaps forms the basis of the psychology of music, as can be interpreted from the predominance of structurally congruent relationships schematised in Deutsch (1982) and Krumhansl (1990).

butterfly schema, and shows that concrete features are not necessary for a formal model. It is proposed that the levels of abstraction used in features of the preliminary model of the butterfly schema are necessary to enable MC in the Classical grammar. The features that form are abstract enough to consistently interact, but not too abstract that they might are meaninglessly generic.

The abstract uniparametrically congruent features of the butterfly schema are distinct forms of chord progression, textural grouping, and harmonic rhythm, in the context of highly abstract tonal and metrical structures. It is posited that it can be intuitively observed that features of the butterfly schema are uniparametrically and multiparametrically congruent. The chord progression is stable, using diatonic chords; the texture has regular correspondence between elements; and the harmonic rhythm is simple and regular. Section 4.2 shows how these uniparametrically congruent features are combined multiparametrically congruently, and that MC is necessary to explain how cognition parses the implicit structure of schemata. However, as pointed out above, a quantitative measure of the UC of each feature is required to support the notion that these features of the butterfly schema are uniparametrically congruent. (Section 5.1 presents such calculations and explains how the abstract uniparametrically congruent features in the Classical grammar and butterfly schema are differently formed from those of the Baroque and Romantic grammars. A statistical survey in Section 5.2 is carried out to provide evidential support that the uniparametrically congruent features of the butterfly schema are consistently multiparametrically congruent and particular in the Classical period, following the rule of MC.)

A main conceptual principle explored in this dissertation is that inductive and associative-statistical theories of musical structure are not as revelatory about cognition as are the present concepts of a tendency for congruence and HS. Many of Gjerdingen's (1988, 2007) schema prototypes can be interpreted in terms of their implicit multiparametrically congruent structure. The preliminary model of the butterfly schema describes the multiparametrically congruent structure of voice-leading schemata. Section 4.3 of this chapter uses the preliminary multiparametrically congruent model of the butterfly schema to examine associative-statistical analyses of voice-leading schemata in Gjerdingen (1988, 2007). The prioritisation of voice-leading features in Gjerdingen's analyses and the consideration of perceptual prototypicality do not uncover the appropriate levels of abstraction as do butterfly schema features; they do not capture the implicit multiparametrically congruent relationships between features. The chapter finds that in Gjerdingen (1988), schemata that are minimally multiparametrically congruent or noncongruent are sometimes deemed valid and used in the statistical survey, whereas highly multiparametrically congruent schemata are often not recognised. By contrast, the butterfly schema shows the implicit MC and multiparametric noncongruence in those structures, which is an alternative, and arguably more powerful explanation of schematic and grammatical structure.

4.1 Determining Multiparametrically Congruent Features

This section contends that highly abstract tonal and metrical structures are perceived and conceived in real-time listening (inside the head) through the bottom-up perception and cognition of less abstract surface features, such as chord progression, texture, and harmonic rhythm. However, tonal and metrical structures are generally caused in musical structure (outside the head) by the tendency for congruence and the transmission of these features in culture, which then govern the formation of less abstract features (chord progression, texture, harmonic rhythm). This can be explained through top-down causation of culture (a full explanation of which is given in Section 5.3 in the top-down hierarchical selection (HS) model). Variability in the tonal and metrical structures of grammars is explored, showing that diverse tonal and metrical structures give rise to diverse chord progressions, textures, and harmonic rhythms. The tendency for congruence governs all features of grammars, both the highly abstract tonal and metrical structures, and the less abstract surface features of grammars.

4.1.1 Defining Universal and Particular Abstract Structures

Of the various features in grammars, tonal and metrical structure are perhaps the most difficult to model because they are the most abstract. While in many psychological studies listeners often report a general appreciation of the notion of key or metre (e.g., Krumhansl

(1990)), and pitch and rhythm have been extensively theorised (e.g., Rameau (1722/1971), Schenker (1935/1979), Lerdahl and Jackendoff (1983)), there is no formula for defining these notions, or general agreement about how they might be verified. Notwithstanding, the highly abstract features of tonal and metrical structure have been broadly explained in various contexts as the conglomeration of less abstract features, such as chords or phrase grouping (e.g., Rameau (1722/1971), Lerdahl and Jackendoff (1983), Krumhansl (1990), Swain (2002), Kaiser (2007b)).

As discussed in Chapter 1, the demarcation between universal and particular structure is not easily drawn. The Gestalt principles of perception (Wertheimer, 1923) and the cognitive constraints on basic pitch spaces (Lerdahl, 2001, pp. 268–274) are the primary universal constraints in many generative theories, which are posited to be reducible to the tendency for congruence (see Section 1.3.2). The Gestalt principles of similarity and proximity mainly govern the form of grouping and metrical structure, and the pitch space constraints limit the distances between pitches, chords, and keys. The universal constraints on basic spaces in Lerdahl (2001, pp. 268–274) condition the types of basic spaces that emerge in grammars (such as the diatonic, octatonic, and hexatonic spaces). Basic spaces can be viewed as psychoacoustic frames that classify the possible ways to cognise pitch space (Section 2.3.7). Universal human cognition seemingly regenerates (*wiedererzeugt* in the sense used by Humboldt (1836, 93)) similar basic spaces in minds. However, while basic spaces are universally constrained (by the tendency for congruence), they are culturally transmitted in particular times and places. So, the employment of particular spaces in a grammar depends on both cultural selection and the tendency for congruence. The same approach can be used to explain metrical structures, which are universally constrained, yet are also culturally selected in particular grammars.

Determining the interacting features in grammars can be problematic because certain features can correlate in a particular piece or situation and not be consistently congruent (as required by the application of the rule of MC). It is thus important to distinguish which features are consistently multiparametricaly congruent as opposed to being merely incidentally congruent (as is done quantitatively in the statistical survey in Section 5.2). Also, it is necessary to show that features are abstract enough to permit consistent MC in the Classical grammar but not too abstract that they are meaninglessly generic. By contrast, some music theories (e.g., Gjerdingen (1988), Byros (2009a)) do not use broad enough categories to permit the establishment of multiparametrically congruent features.

4.1.2 Bottom-up and Top-Down Causation

The concept of causal levels must be introduced to explain metrical and tonal structure. The causes of these highly abstract features are of at least two kinds, which can be explained using the concept of bottom-up and top-down causation (Ellis, 2005, 2008; Okasha, 2012) (fully explained in Section 5.3.2). The bottom-up, real-time perceptual-cognitive cause of metrical and tonal structure, which occurs inside the head, differs from the top-down cultural, historical, and cognitive causes of these structures, which occur outside the head. In real-time listening, the concrete, less abstract surface features (such as chord progression, textural grouping, and harmonic rhythm) are generally perceptually and cognitively organised into the more abstract tonal and metrical structures (as shown in a hierarchy of causation in listening in Figure 4.1). This perceptual-cognitive ordering (of concrete, and less abstract features into collections of highly abstract features), such as where rhythmic structure is organised into metrical structure, or where pitch structure is organised into chords and thereafter into tonal structure, is generally assumed in music theory (e.g., Piston (1941), Swain (2002)) and music psychology (Deutsch, 1984; Krumhansl, 1990). Such abstraction is possible because listeners intuitively understand the implicit multiparametrically congruent or noncongruent relationships, and so can categorise the structure accordingly (as discussed in Section 1.3).[24]

[24] Listeners do not prefer MC in real-time listening but become intuitively aware of multiparametrically congruent and noncongruent relationships, which enables them to categorise multiparametrically congruent structure (as argued in Section 1.3.2). By contrast, the tendency for congruence (measurable through the rule of MC) is part of the top-down historical cultural and cognitive cause of butterfly schemata, as shown in the top-down HS model.

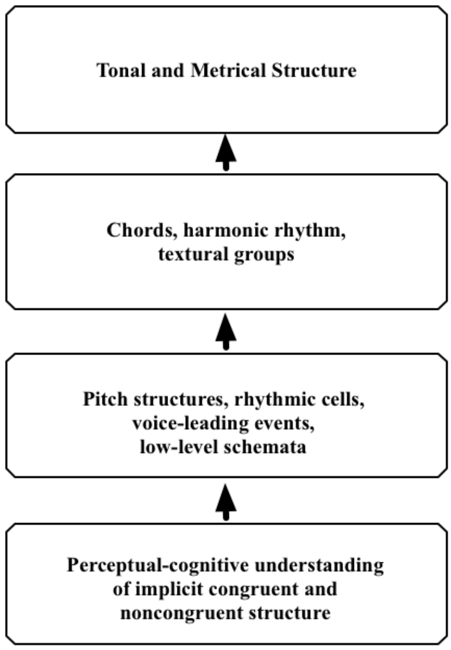


Figure 4.1 Real-time, perceptual-cognitive, bottom-up hierarchical cause of musical structure.[25]

As mentioned, the top-down, cultural, historical, and cognitive cause (the tendency for congruence) of these structures reveals a reverse order of causation to the bottom-up, real-time cause, shown in Figure 4.1. Indeed, in the top-down cause, highly abstract metrical and tonal structures are culturally selected and transmitted in grammars, as in the Classical grammar, while constrained by the tendency for congruence. Cone (1968) has shown the prioritisation of particular levels of metrical hierarchy in certain periods and styles, such as in the Baroque, Classical, and Romantic periods (discussed in Section 4.1.4). In history and culture, lower-level features, such as chord progression, harmonic rhythm, and textural grouping are selected by these highly abstract upper-level features, through top-down causation. That is, culture, history, and cognition selects and transmits metrical and tonal

[25] The model of bottom-up, real-time listening in Figure 4.1 is a broad simplification of how listening actually takes place. It is a possibility that real-time listening involves the interaction between bottom-up and top-down causation, as causation has been proposed to occur in the real world in Ellis (2005, 2008). This type of multidirectional causation might be challenging to model because it involves a consideration of various emergent, non-algorithmic processes.

types, governing the formation of less abstract surface features that form through top-down causation. When listeners hear musical structure in real-time, they unravel the structure in the converse order, through intuitive, *a priori* perceptual-cognitive bottom-up causation, constructing the implicit abstract tonal and metrical structure from the concrete and less abstract surface features (as shown in Figure 4.1).

It is probable that the real-time perceptual-cognitive cause of metrical and tonal structures applies to every instance of real-time listening. While selection in certain cultures can necessarily be top-down, selection in cognition is never necessarily top-down, because cognition is intuitive. Thus, regardless of the particular arrangement of features in certain grammars, music is so diverse that the benefits for listeners in applying learned schemata are few. A consideration of the top-down cultural cause of butterfly schema permits the formalisation of the multiparametrically congruent and top-down HS models of the butterfly schema.

4.1.3 Tonal Structure

Schenkerian theory (Schenker, 1935/1979), *GTTM* (Lerdahl and Jackendoff, 1983), and *TPS* (Lerdahl, 2001) generally consider tonal structure (and chord progression and harmonic rhythm generally) to be an 'output' of their generative grammars, experienced as prolongational tension. Therefore the influence of harmony and prolongation on other features, such as grouping or metrical structure is not fully considered. Mirka (2009, p. 50) argues that the lack of specific well-formedness rules and preference rules in *GTTM* for the accent-producing factor of harmony and harmonic rhythm is puzzling because they are often significant influences on other features. While chord progression and harmonic rhythm are broadly included in the various accentuation rules (Lerdahl and Jackendoff, 1983, pp. 76–89), these are not detailed enough to account for the specific effect of harmony and harmonic rhythm in some grammars, such as the Classical grammar, where these are integral multiparametrically congruent features.

The presentation of harmony as a cognitive experience in *GTTM* is typical of the general focus on cognitive experience and modelling in generative theories. However, generative theories generally do not fully account for the effect of cognition on musical structure. While tonic harmony prolongation (or more generally, tonal structure) is a fundamental feature of many grammars, it might be structured in a variety of ways in those grammars. For example, tonic chord prolongation in the Classical grammar is relatively stable globally, but in the Baroque contrapuntal grammar (as, for instance, in the music of J.S. Bach) tonal structure is characterised by alternating tonal centres in local contexts. Moreover, chord progression is more significant for creating MC in the Classical grammar than many other grammars. Broadly, in the contrapuntal grammar of the Renaissance, chord progression is often a consequence of voice-leading interaction (Berry, 1976; Swain, 2002). In general terms, the Baroque grammar — which sits between the Renaissance and Classical periods — is an admixture of the two, using chord progressions that are sometimes a consequent of, and at other times an antecedent to, voice-leading considerations. Tonal structure is dealt with more comprehensively in Section 5.1, where a discussion of the interaction between the universal and particular continues, but using the more precise conception of the pitch-space theory (of Lerdahl (2001)) to formalise chord progressions.

4.1.4 Metrical Structure

Figure 4.2 shows a hierarchical grid-like representation of metrical structure. The note values are not absolute, intending only to illustrate the regular, duple, and hierarchical relationships between levels of metrical structure. In the Classical grammar, metrical structure is generally regular, duple, and hierarchical at the level of regular functional harmonic change (often the 5^{th} level) and at two immediately higher levels (often the 6^{th} and 7^{th} levels). (This structure breaks down at levels higher and lower than these.) This universally constrained type of metrical structure is culturally transmitted during the period. This means, through top-down HS, the more abstract surface features of metrical and tonal structure ultimately give rise to the less abstract chord progression, harmonic rhythm, and textural grouping in the Classical grammar (although in real-time listening metrical structure is constructed in the converse order). This order of dependency between features can vary between grammars and contexts.

In the contrapuntal grammars of the Renaissance and the Baroque, harmonic rhythm is less regular, duple, and hierarchical at the 5th, 6th, and 7th levels, and is therefore not necessary for creating multiparametrically congruent relationships with other features (further elaborated in Section 5.1.2.1). In art music of the Romantic period (such as that of Wagner of Brahms), harmonic change is also less regularly congruent with metrical structure at these levels, and so is also not significant for generating multiparametrically congruent relationships.

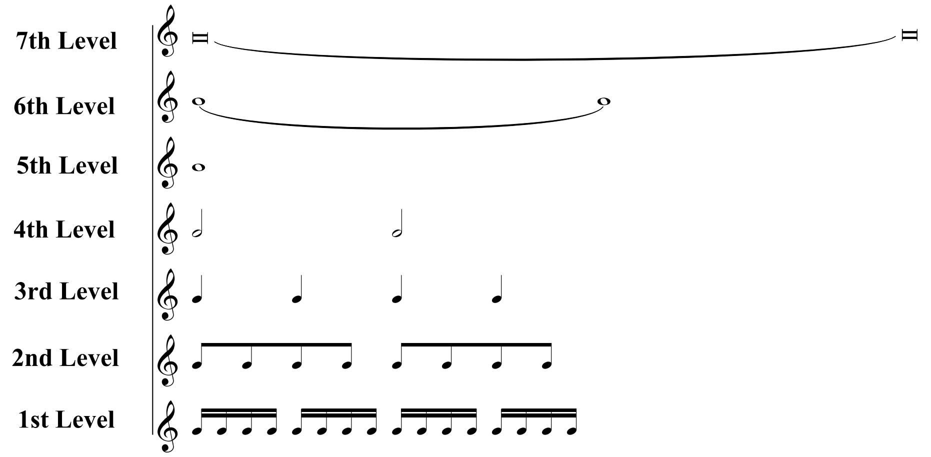


Figure 4.2 Hierarchical levels of division in metrical structure.

The present notion that universally constrained metrical and tonal structures manifest in particular ways in grammars has general support in music theory. However, metrical structure is generally considered from three perspectives. It is either the outcome of an implicit universal mental capacity (Lerdahl and Jackendoff, 1983; Temperley, 2001; London, 2004), the product of emergent structuring (Hasty, 1997), or a cultural construct (Cone, 1968; Berry, 1976; Lester, 1986; Benjamin, 1984; Kramer, 1988; Rothstein, 1989). The first conception is given almost unanimous support in music psychology (e.g., Huron (2006)), but is popular in music theory also (e.g., London, 2004). London (2004) argues that metre is generated through the psychological entrainment to regularity. Listeners are proposed to necessarily construct an invariant psychological schema, or grid, in response to such metrical cues (Lerdahl and Jackendoff, 1983; Temperley, 2001; London, 2004).

The second and third views of metrical structure, as emergent or cultural types, suggest that metrical structure is flexible and can have degrees of variability. In Hasty (1997), rather than being invariant and universal, metrical structure is formed through mental projections that are realised through the emerging surface structure — and so metrical structure is not cognitively implicit. This, a nominalist view of metrical structure, therefore suggests metrical structure is dependent on the interaction of emergent concrete features. The varied types of metrical structures that form in Western music support this view. The establishment of metrical structure is conditional on culture, genre, style, and the particular composer or piece (Hasty, 1997). This perspective contends that metrical structure might appear to have an implicit or essential structure because it is normatively regular, duple, and hierarchical, but that this is not necessarily so, since it can vary if cued by irregular and non-hierarchical surface features (Hasty, 1997, pp. 168–169).

Returning to the first perspective described above, metrical structure is often modelled as a cyclical well-formed structure (e.g., London (2004)). Zuckerkandl (1956, pp. 169–200) similarly describes metre as a fixed wave-like phenomenon. However, it is problematic to use a general model to define an abstract and variable concept. The wave-like conception does not encompass many other characteristics of metrical structure, such as its hierarchical structure (as described in Lerdahl and Jackendoff, 1983), and its frequent irregularity (Benjamin, 1984). Metrical structure is also conceived as a grid, which is a well-formed, hierarchical structure. This is not completely appropriate because metrical structure is not often well-formed in actuality. Benjamin (1984, pp. 371–376) and Kramer (1988, p. 102) posit that metrical structure is usually, but not necessarily, hierarchical. Kramer (1988, p. 102) contends that 'metrical structure is deeply hierarchic, but grids can exist on several levels, simultaneously, [and] we are quite capable of understanding irregularities that are subsumed into deeper-level regularities'.

Understanding metrical structure through the strictures of well-formedness is a limitation for modelling the diversity of musical structures (Benjamin, 1984). When applied too rigidly, well-formed grids are idealist fictions, limiting understanding of the intrinsic variety found in metrical structures. Even though duple, regular, and hierarchical metrical structure is culturally transmitted and manifest in notation of the Classical period, in actuality metrical structure does not adhere to well-formedness, since the regular, duple hierarchy can be controverted at various levels. Despite views to the contrary (cf. Lerdahl and Jackendoff, 1983), ill-formed metrical structures are possible and consistently occur in grammars. For example, grouping structure (which cues metrical structure) can be frequently out of phase (or 'dissonant') with the dominant metrical structure, as shown in Rothstein (1995), suggesting metrical plurality. Mirka (2009) also tacitly reveals the limitations of metrical grid schemata, identifying metrical dissonances in pieces that occur through conflict between the emerging grouping structures and established metrical structures.

As discussed, London (2004) presents the strongest argument for the notion of well-formedness, positing that listeners are only able to hear a passage in terms of a single metrical schema, entraining to a single regular input signal. That is, London (2004, pp. 79–86) sees the psychological process of attending to metre as negating the possibility of hearing two or more metres simultaneously. This view diminishes the inherent metrical complexity and variability in metrical structure and flatly contradicts the notion of polymetre. In the grammars of the Renaissance and Baroque, grid-like structure occurs reliably only at lower levels, whereas upper levels are much freer. Moreover, while grid-like schemata are common at certain levels of metrical structure in notated Western musics, they are not so common in non-Western aural cultures (Benjamin, 1984, p. 358). Temperley (2001, pp. 292–297) asserts that grouping and metre in Western and non-Western musics generally conform to metrical grids, but Temperley (2008, 2009c) tacitly contradicts this point when presenting examples where metrical structure is highly differentiated, showing complex metrical structures and 'hypermetrical transitions' in Classical music.

The general dichotomy between idealist (London, 2004) and nominalist (Hasty, 1997) theories of metrical structure requires synthesis. The model of the butterfly schema (presented in Section 5.1) follows a theory of metrical structure where the grid schema is implicit at certain levels of metrical structure, but not the whole structure, since it is constrained, not fully governed, by the tendency for congruence. The implicit constraint for duple, regular hierarchies is manifest at particular levels of metrical hierarchy, but admits variably congruent structure. Even in the most regular, duple, and hierarchical Western grammars, the grid schema only occurs at particular levels — it is not absolute. Thus grid-like structure is by degree, and must be employed flexibly, rather than used as an invariable rule.

The discussion so far makes the case that models of tonal and metrical structures in particular grammars require a more fine-grained approach than given in generative theories. Greater particularity in models also requires considering abstract surface features (such as chord progression, texture, and harmonic-rhythm) in a more coherent way. Lerdahl and Jackendoff (1983, pp. 75–79) posit that melodic phrase grouping dictates the inference of metrical structure. Likewise, grouping and metre are interdependent in Cooper and Meyer (1960). However, Hasty (1997) posits that metre is a construct that is actually inseparable from rhythm. Other theories make more marked distinctions between these two features (Lester, 1986; Kramer, 1988; Rothstein, 1989). Rothstein (1989, p. 156) and Mirka (2009) posit that while there is generally a congruent relationship between grouping and metre in Classical music, there are also many examples of noncongruent metrical structure, characterised by metrical deletion (which Rothstein (1989, p. 52) terms 'reinterpretation') and phrasal elision. (Metrical deletion is when metrical structures are perceived to overlap, and phrasal elision is when grouping structures overlap (Rothstein, 1989, p. 52).) Significantly, metrical deletion can occur without phrase overlap, but phrase overlap cannot occur without metrical deletion (Rothstein, 1989, p. 52), demonstrating that metrical structure is a more abstract and fundamental aspect of grammars than melodic phrase grouping. Metrical structure forms from a number of interacting features, which are not easily controverted. As discussed, *GTTM* proposes that metrical structure is primarily inferred through melodic phrase grouping (Lerdahl and Jackendoff, 1983, pp. 75–79). While other preference rules influence metre in *GTTM*, they are presumably of less importance. By contrast, in the views of other theorists (Lester, 1986; Kramer, 1988; Rothstein, 1995, 1989; Swain, 2002; Mirka, 2009), a variety of features are proposed to interact to produce metrical structure.

Cone (1968) proposes models of various types of metrical structures situated in the Baroque, Classical, and Romantic periods. Cone (1968) suggests that the underlying 'metrical units' of these periods differ in terms of the level of the metrical hierarchy that is prioritised. The metrical unit is the level of metrical structure at which there is greatest combined structural accentuation (presumably between the more concrete features, such as chord progression, texture, and harmonic rhythm). 'Textural accent' is also important for the inference of metrical structure, which is the accentuation engendered through the combined onsets of voices in texture and extremes of registers (Lester, 1986, pp. 29–33). However, textural accent works in conjunction with other features to inform metrical structure.

In the butterfly schema model, the concept of 'textural grouping' is introduced, which is the regular, duple, and hierarchical grouping of textural accents that informs metrical structure and creates MC. However, texture is given little significance for the inference of metrical structure in *GTTM*, where it is incorporated as an appendage to grouping, through the concept of 'fusion' (Lerdahl and Jackendoff, 1983, pp. 153–155). Texture is given more attention in Lerdahl (2001) and Temperley (2001), where it is conceived through the notion of contrapuntal grouping. Textural grouping contrasts with these ideas because it prioritises the grouping of texture at particular levels of metrical structure, while grouping is resisted at other levels, depending on the grammar (see Section 5.1.2).

Cone (1968, p. 66) argues that the dominant metrical level (or unit) in Baroque music is the beat (generally the 2^{nd} and 3^{rd} level of metrical structure, shown in Figure 4.2). However, in Classical music, the bar level (generally the 5^{th} level) is the main metrical unit (Cone, 1968, p. 72). Although the beat level is not insignificant in Classical music, it is not as regularly accented or prioritised as it is in Baroque music. In Classical music the beat level can often be irregular during the course of a piece. That is, in Classical music the notated bar has variable subdivisions (Cone, 1968, p. 72). Lester (1986, pp. 127–156) confirms these metrical differences between Baroque and Classical music, pointing out that freer, less grid-like structures emerge at higher levels in Baroque music, above the bar level (generally the 5^{th} level of metre). The hypermetrical freedom in some Baroque music (particularly that of J.S. Bach) works complementarily to the yoking of the musical action to the beat level. By contrast, bar-level and hypermetrical freedom is less common in Classical and Romantic music (Cone, 1968, p. 74, p. 79).

Romantic music requires more cautious analysis because it contains a broader collection of sub-grammars with a greater variety of metrical profiles. Both Cone (1968, p. 79) and Lester (1986) posit that metrical structure in the Romantic period becomes more regular, or more grid-like during the latter half of the nineteenth century, as seen in the music of composers such as Bruckner and Offenbach. However, the music of late Wagner, for instance, does not fit this mould. In late Wagner, textural grouping structure, and therefore also metrical structure, is often irregular (Rothstein, 1990, pp. 276–305). Thus, in defining metrical structures of this period, corpora must be carefully delimited to account for the multiparametrically congruent structure.

In the Classical grammar, and, by extension, the butterfly schema, metrical structure is subject to universal cognitive constraints (reducible to the tendency for congruence), and particular culturally conditioned constraints, exhibiting well-formed or grid-like tendencies at particular levels of hierarchy. That is, the metrical structure of the butterfly schema and Classical grammar is a particular manifestation of the tendency for congruence. A number of theorists support the view that chord progression and harmonic rhythm are regularly congruent with textural grouping and metrical structure in the Classical style — although attributing such interaction to tonal music generally. For example, Rothstein (1989, 22), Lester (1986, p. 58, pp. 66–68, p. 159), Swain (2002), and Mirka (2009, p. 50) posit that harmonic change strongly influences metrical structure. Lester (1986, p. 66) also argues that harmony is decisive for establishing metre, pointing out that '[h]armonic change is the single most powerful meter-producing factor'. Harmonic change, or harmonic rhythm, is also proposed to be a cue for phrase structure in Rothstein (1989, pp. 21–22) (which is itself a cue for metrical structure in *GTTM*). Rothstein's (1995, p. 173) 'rule of harmony' posits that changes of harmony preferably occur at the inception of strong beats in metrical structure.

4.2 Multiparametric Congruence in the Classical Grammar and Butterfly Schema

This section shows that multiparametric congruence (MC) characterises the implicit structure in and between the features of the butterfly schema and the Classical grammar. MC is intuitively understood in musical structure in real-time (as argued in Section 4.1.2) and is necessary for the identification of the butterfly schema. In Section 4.2.2, a number of concrete localised schemata in the Classical grammar are examined that are similar to the concrete voice-leading schemata of Gjerdingen (1988, 2007). It is argued that abstract multiparametrically congruent features form necessary components of schemata in the Classical grammar, whereas concrete schemata do not. These abstract schemata are definable according to the multiparametrically congruent relationships between their features.

Indeed, the present conception of MC provides a more powerful system for defining the schemata than the use of congruence in many generative theories of music because it permits more flexible readings of the interaction of features, taking into account the particular manifestation of MC, such as those of the Classical grammar and the butterfly schema. In Section 4.2.3, a preliminary version of the butterfly schema is presented. In the butterfly schema, the abstract surface features — chord progression, textural grouping, and harmonic rhythm — and the highly abstract features (tonal and metrical structure) are multiparametrically congruent.

4.2.1 Combining the Features of the Butterfly Schema Through Multiparametric Congruence

GTTM and *TPS* tacitly incorporate the tendency for congruence through well-formedness rules and preference rules, although MC occurs in more varied ways than described in these theories. In *GTTM*, it is assumed that congruence binds the four components of *GTTM*: grouping structure, metrical structure, time-span reduction, and prolongational reduction. However, the rules of *GTTM* can conflict with the implicit congruent structuring found in specific Western and non-Western grammars.

This can be shown through considering various abstract multiparametrically congruent relationships in grammars. Kaiser (2007b) employs the notion of an 'abstract schema' to refer to a class of patterns (such as a tonal system, or harmonic system, etc.) which have elements that are not instantiated individually at any one time, but which are concrete tokens that infer that class or abstract structure. Using the example from Kaiser (2007b), tonal musical pieces are concrete instantiations of the abstract schema of the tonal idiom, while the class of structures that belong to the tonal idiom is not limited to any particular tonal pieces. Likewise, chord progression, textural grouping, harmonic rhythm, and tonal and metrical structure are broad-stroke abstract features because they can comprise various instantiations of concrete features. (When these features are considered individually they are uniparametrically congruent, and when combined they are multiparametrically congruent.) MC is abductively understood through rational cognition in real-time. Humans are able to comprehend the implicit MC and multiparametric noncongruence in structure. Grammars and local schemata have features at certain levels of abstraction that are consistently multiparametrically congruent, but vary depending on the particular structure. In this respect, many generative theories do not provide generalisable rules that could yield coherent readings of the interaction between congruent features of various grammars. *GTTM* uses well-formedness and preference rules that are too specific to apply universally. Moreover, the features defined in *GTTM* are often not at the right level of abstraction appropriate for particular grammars.

Figure 4.3 shows how well-formed metrical grids in Western music tend to be formed from regular binary grouping structures. The universal constraint for regular, binary, and hierarchical metrical structure is strongly reinforced at particular levels, depending on the grammar. However, various structures can feasibly form at other levels, shown in Figure 4.3, where the 2nd level of metrical structure has ternary groupings.

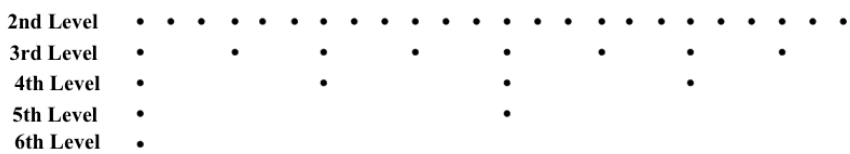


Figure 4.3 A metrical grid (M) based on Metrical Well-Formedness grids.

Textural grouping is an abstract uniparametrically congruent feature of the Classical grammar and butterfly schema, necessary for MC. It is inferred through correspondences between note onsets in different voices of texture, generating the implicit metrical grid at the level of regular functional harmonic change (often the 5th level) and at two immediately higher levels in the Classical grammar and butterfly schema. It interacts with other multiparametrically congruent features to cue metrical structure at these levels. This contrasts with how metrical structure is generated in *GTTM*, where melodic phrase grouping is the primary influence on metrical structure. The theory of textural grouping also contrasts with the contrapuntal grouping components of Lerdahl (2001) and Temperley (2001) because it works at specific levels of metrical hierarchy, depending on the grammar. More pointedly, textural grouping in the Classical grammar regularly corresponds with other parametric features to produce metrical structure at particular levels (at the level of regular functional harmonic change and at two immediately higher levels). Textural grouping is an important feature in the Classical grammar at these levels because it more regularly corresponds with metrical structure than melodic phrase grouping. Figure 4.4 shows the correspondence between the hierarchy of textural grouping and metrical structure in the Classical grammar.

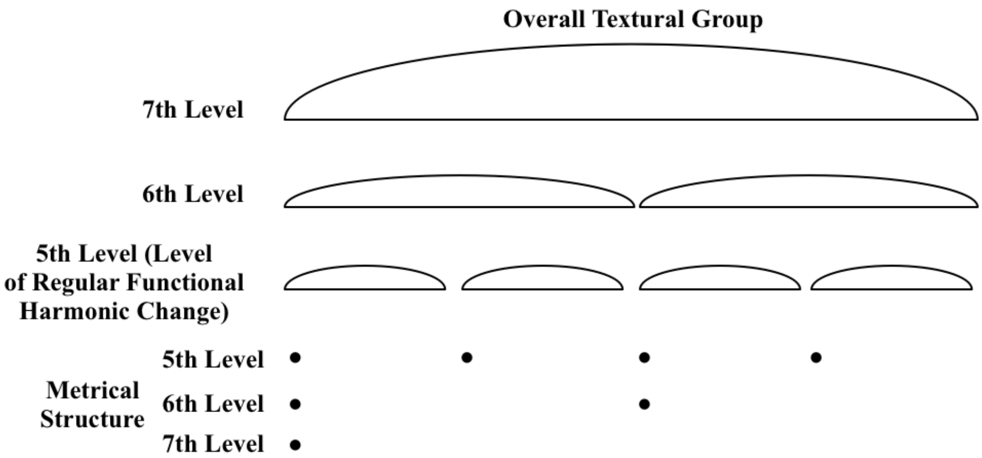


Figure 4.4 Correspondence between textural grouping and metrical structure at the level of functional harmonic change and at two immediately higher levels in the Classical grammar.

In Figure 4.5, a parent schema, P, comprises abstract uniparametrically congruent features that interact multiparametrically congruently (however, each feature — chord progression, textural grouping, and metrical structure — can have various concrete instantiations). The multiparametrically congruent interaction between these features is ubiquitous throughout the Classical grammar. The chord progression of P is a chord V in the first bar to a chord I in the second bar. The textural grouping occurs at the level of regular functional harmonic change and at the level immediately above. The harmonic rhythm changes with the bar (which might be described as ‘metrical rhythm’). The chord progression of the schema is governed by tonic harmonic prolongation. The hemispheres represent the textural grouping structure, the dot diagram (from *GTTM*) represents the metrical structure, and chords V and I depict the (relatively) abstract chord progression. In sum, these less abstract surface features are multiparametrically congruent with the highly abstract metrical and tonal structure (which is tonic harmony prolongation).

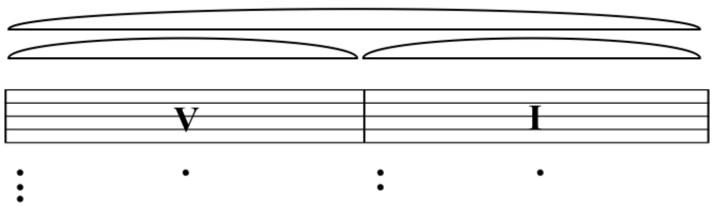
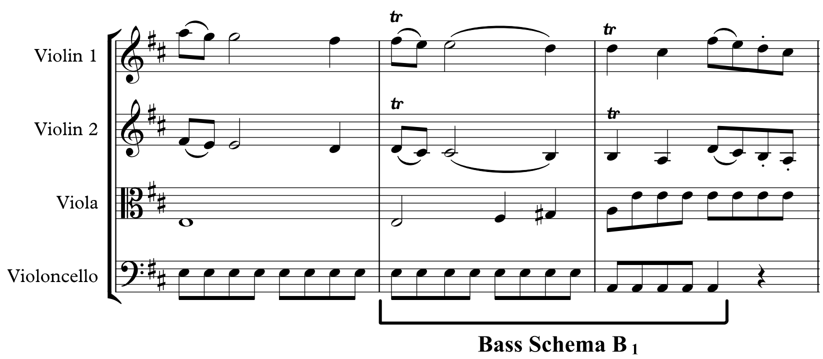


Figure 4.5 Abstract parent schema (P).

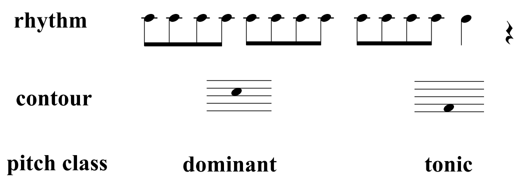
The P schema is broadly commensurate with the second section of a butterfly schema, the basic form of *GTTM* (Lerdahl and Jackendoff, 1983, pp. 188–189), the ‘archetypal phrase features’ of *GTTM* (Lerdahl and Jackendoff, 1983, p. 289), and also the final half of many binary voice-leading schemata of Gjerdingen (1988, 2007). Also, P is an abstract parent schema that can comprises many multiparametrically congruent child sub-schemata — such as the second section of many of the voice-leading schemata of Gjerdingen (see Section 4.3.1). There are probably a large number of concrete instances of P in Western music, while it is more common in some grammars than others. For example, it presumably appears in instrumental music of the Classical period more often than in instrumental music of the Renaissance or Baroque periods.

4.2.2 A Limitation of Concrete Schemata

Figure 4.6a shows a concrete bass schema, B_1, which comprises a network of concrete features, and implies a number of abstract features. The network of concrete features of B_1 is shown in Figure 4.6b. B_1 contains the concrete elements of pitch class (dominant–tonic), contour (a falling 5th), and rhythm (quavers followed by a crotchet). B_1 has the implicit abstract features of schema P (shown in Figure 4.5) (although P only denotes abstract features, it does not specify concrete features). The falling 5th and quaver rhythm can be subsumed within the abstract textural grouping feature of P, and the concrete pitch class (dominant–tonic) can also be generalised into the abstract chord progression of P. Therefore, listeners can intuitively understand the implicit abstract structure of P upon hearing B_1. However, the converse is not true: listeners do not necessarily hear B_1 upon hearing P.



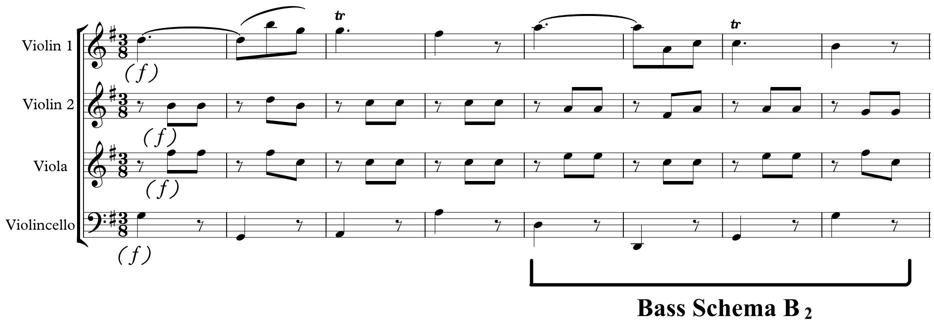
a.



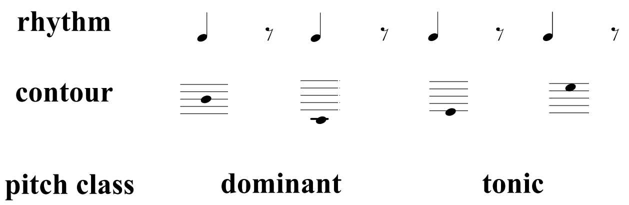
b.

Figure 4.6 Bass schema, B_1, in Mozart String Quartet No. 2 in D Major, K. 155 (1772), i, bars 14–16, score (a), network analysis of B_1 (b).

Figure 4.7a–b depicts a bass schema, B_2, which has concrete features that mostly do not match those of B_1, although B_1 and B_2 have the same implicit structure of P. B_2 has single crotchet notes that differ from the quavers of schema B_1, and observes a different contour, rising instead of falling from the dominant to tonic note. Thus while B_1 and B_2 conform to the abstract class, P, they have differing concrete features, excepting pitch class (they both involve movement from the dominant to tonic pitches) (shown in the network structures of Figures 4.6b and 4.7b).



a.



b.

Figure 4.7 Bass schema, B_2, in Mozart, String Quartet No. 3 in G major, K. 156 (1772), i, bars 1–8, score (a), network analysis of B_2 (b).

If the common feature between B_1 and B_2, dominant–tonic pitch class, were included in P it would give this abstract schema much more specificity, limiting its flexibility and also changing its status from an abstract schema into a concrete schema, which is shown in schema G in Figure 4.8. G resembles the voice-leading schemata of Gjerdingen (1988, 2007)[26] since it assigns concrete pitch classes to abstract congruent schemata. That is, G is effectively the same schema as P, but with the added concrete feature of dominant–tonic pitch classes, which are also used in B_1 and B_2. This specialised pitch-class feature in the bass makes G more specific, but also more exclusive than P. This limits its applicability because it can only be evoked in specific instances where the pitch-class structure in the bass is present.

[26] Gjerdingen (1988, 2007) does not explicitly define this particular schema, G, but such attention to voice-leading is axiomatic in his application of schema theory.

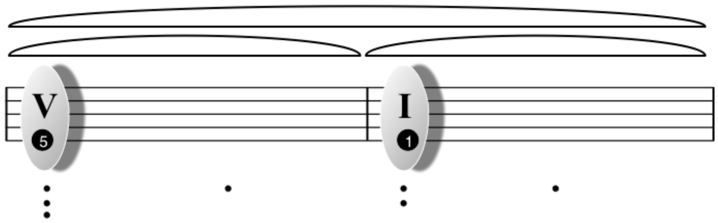


Figure 4.8 Schema G.

Indeed, this mixture of abstract and concrete structures in the Gjerdingenian schema, G, means it is a fundamentally different type of schema to P because the latter comprises only abstract multiparametrically congruent structure. Since, the bass pitch class in G can be subsumed under the implicit abstract order of P, it might be viewed as an instantiation of the implicit multiparametrically congruent structure in P. In this sense, G is sufficient but not necessary, to invoke the abstract multiparametrically congruent structure of P. While the features of P can be flexibly reified, the more concrete schema, G, is a distinct schema that is less flexible because it cannot schematise the possible types of concrete multiparametric interaction possible under P.

4.2.3 A Preliminary Model of the Multiparametrically Congruent Butterfly Schema

As discussed in Section 4.1, the portrayal of the butterfly schema contrasts to a degree with the implementation of grammatical rules in *GTTM* because the latter reifies multiparametrically congruent relationships that are often more diversely organised in grammars. *GTTM* conveys a particular serial order of interaction between rules and components (from grouping structure through to metrical structure, time-span segmentation, and prolongational reduction, respectively) that does not pertain to the interaction of features in the butterfly schema and Classical grammar. This conflict with *GTTM* will be examined through consideration of the preliminary version of the multiparametrically congruent butterfly schema (see Figure 4.9). (This preliminary version is a simplification of what will be presented in Section 5.1.4.).

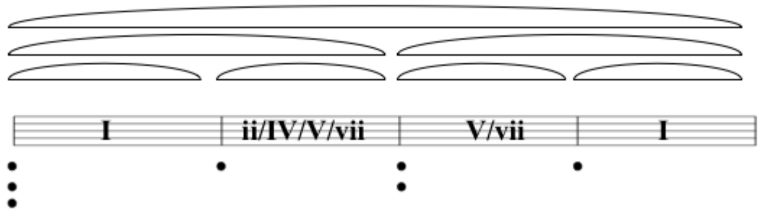


Figure 4.9 Preliminary version of the butterfly schema.

The whole schema is governed by tonic harmony prolongation. The chord progression uses uniparametrically congruent chords, which are close to the tonic in pitch space. Broadly, tonic chords are more congruent than dominant chords, and both are more congruent than supertonic, subdominant, and leading-note chords. Indeed, tonic and dominant chords tend to occur at strong points of the textural and metrical structure. The harmonic rhythm generally changes with the bar. The textural grouping, at the level of regular functional harmonic change and at two immediately higher levels, corresponds with the metrical structure.

The opening of Mozart's *Ave verum corpus*, K. 618 in Figure 4.10, exhibits an abstract multiparametrically congruent butterfly schema. An interpretation of this passage in terms of the rule system of *GTTM* does not fully account for how the implicit congruent structure of the passage is understood in real-time listening, also, the order of dependency between features propounded in *GTTM* does not show how the passage is generated through the particular MC of the Classical grammar.

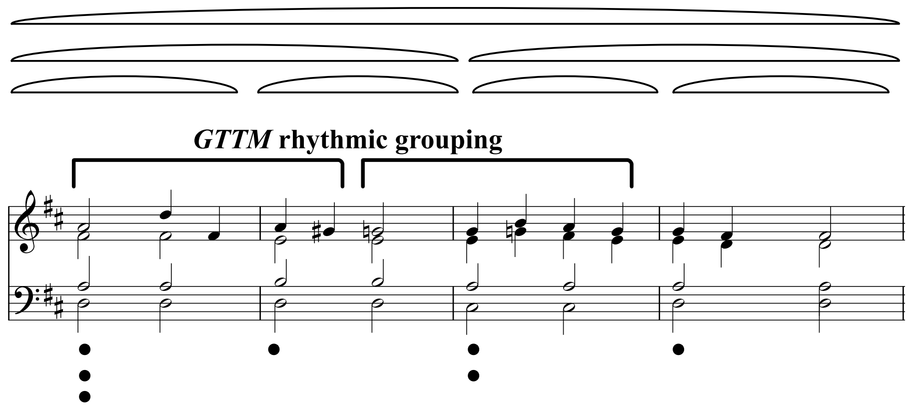


Figure 4.10 Mozart, K. 618 (1791), bars 1– 4.

In *GTTM*, the order of dependency between features (or components of the grammar) is broadly: grouping structure, metrical structure, time-span reduction, and prolongational reduction, respectively. The grouping of a melody is viewed as fundamental for the inference of metrical structure. However, in Figure 4.10, the application of GPR 6 (Parallelism) and MPR 1 (Parallelism) does not permit the real-time inference of metrical structure. Also, while GPR 2 (Proximity) might establish a boundary between bars 2 and 3, this alone cannot establish the multiparametrically congruent grammatical structure. Bars 2 and 4 have the same rhythm, but bars 1 and 3 are not similar in terms of rhythm or pitch. Although GPR 6 (Parallelism) refers to 'parts of groups' (Lerdahl and Jackendoff, 1983, p. 51) it is questionable whether bars 1 and 3 can inform the metrical structure because they are not parallel in terms of rhythm or pitch. An application of GPR 5 (Symmetry), which describes the preference for duple structuring, permits the parsing of the grouping and metrical structure. However, GPR 5 (Symmetry) is a similar notion to the universal preference for duple, regular hierarchies incorporated into the butterfly schema (see Section 5.1.2), but is rather a top-down preference for congruence, rather than a bottom-up explanation of parsing in real-time, so is of questionable significance.

Melodic phrase grouping cannot be a primary consideration for the real-time parsing of the grammatical structure of this passage. The grouping of the melody, which is presumed to be a main cue for metrical structure in *GTTM*, is here actually noncongruent with the textural grouping and metrical structure (suggesting that melodic phrase grouping might not be a cue for metrical structure generally in the Classical grammar). The lack of congruence between the grouping of the melody and textural grouping is an occasional characteristic of the Classical grammar. In terms of its rhythm and contour, the grouping of the melody actually cues a 3/2 metrical structure in this instance, conflicting with the actual and notated metrical structure, of 4/4. The textural grouping, which is the corresponding onsets in separate voices, occurs at the level of regular functional harmonic change and at two immediately higher levels. This is more significant in this passage for UC than melodic phrase grouping. Melodic grouping structure is a superficial feature here, which apart from its influence within textural grouping, merely provides a noncongruent nuance.

The multiparametrically congruent grammatical structure is intuitively understood by listeners in real-time primarily through the multiparametrically congruent interaction between the uniparametrically congruent features of chord progression, textural grouping, and harmonic rhythm. *Ave verum corpus* has a congruent chord progression; the chords are close to the tonic in pitch space, with chord changes at strong points of the metrical structure. Indeed, a combination of congruent chord progression, the regular onset of chords on the first beat of each bar (regular harmonic rhythm), and the regular texture, defines the multiparametrically congruent structure of this passage. In bottom-up real-time listening, the more abstract features (of tonic harmony prolongation and regular metrical structure) are generated through the interaction between these less abstract surface features. This order is the converse when considering the historical and cultural cause of this structure, which is the tendency for congruence and cultural selection, which is top-down. By contrast with *GTTM* (which is a cognitive explanation of listeners experienced in the tonal idiom), the causal seriality of parametric interaction in this passage in real-time listening runs from textural grouping, to chord progression, to harmonic rhythm, then to metrical structure and tonal structure. As discussed, *GTTM* mainly considers tonic harmony prolongation primarily as an output of the system, experienced as tension and relaxation patterns in cognition, or as a top-down factor in the time-span reduction and prolongational reduction. However, chord progression, textural grouping, and harmonic rhythm are important initial abstract surface features necessary to invoke the multiparametrically congruent highly abstract metrical and tonal structure in real-time listening.

This section has argued that a particular type of abstract MC is implicit in the butterfly schema and Classical grammar. The abstract features in the Classical grammar, the chord progression, textural grouping, and harmonic rhythm, are important for MC. The butterfly schema combines these grammatical features in a multiparametrically congruent structure. Listeners must intuitively understand how these abstract features form an implicitly multiparametrically congruent structure in real-time because there can be no formula for representing the infinitely diverse multiparametrically congruent (and noncongruent) features that can form in structure. In the next section, the butterfly schema will be shown to correspond with the implicit structure of some of Gjerdingen's (1988, 2007) voice-leading schema prototypes.

The order of dependency between features in the butterfly schema and Classical grammar in real-time listening (and the top-down historical, cultural, and cognitive order of dependency) partly conflicts with that of *GTTM*. Melodic phrase grouping is not as significant as textural grouping and harmonic rhythm for defining the butterfly schema and Classical grammar. As discussed above, in real-time listening, which is a bottom-up cause of musical structure, the abstract surface features (chord progression, harmonic rhythm, and textural grouping) invoke highly abstract features (metrical and tonal structure). However, in the top-down cause of musical structure, which unites the tendency for congruence and cultural selection, a converse order of causation occurs between features, where more abstract features govern less abstract features. Metrical structure and tonal structure are universally constrained by the tendency for congruence, but are particularly manifest — since they are memetically transferred by culture, evinced by the commonality of particular metrical and tonal types in certain cultures.

4.3 Congruent and Noncongruent Schemata

This section shows how MC is a structural characteristic of butterfly schemata in the Classical grammar. Many voice-leading schemata prototypes of Gjerdingen (2007, 1988) can be defined by their abstract implicit multiparametrically congruent structure through representation as child schemata of the parent butterfly schema. Therefore, using the butterfly schema as a template, the voice-leading analyses of Gjerdingen are examined in terms of their multiparametrically congruent grammatical structure. It is argued that since the schema analyses of Gjerdingen prioritise voice-leading events, they do not fully consider the multiparametrically congruent structure. Voice-leading schemata that are minimally grammatically congruent are sometimes validated and included in Gjerdingen's (1988) statistical survey, whereas highly multiparametrically congruent schemata are not validated and not included in his statistical survey. Gjerdingen's methodology contrast with that of the present framework because schemata are here defined as particular, but abstract, multiparametrically congruent structures.

4.3.1 Contrasting Definitions of Schemata

The presentation of schema theory in Gjerdingen (1988, 2007) is attentive to the historical situatedness of schemata. In Gjerdingen (1988), there is a focus on a single schema, the 1–7… 4–3 (Figure 4.11), whereas Gjerdingen (2007) examines a number of Galant schemata. Many of Gjerdingen's (1988, 2007) voice-leading schemata can be depicted as child schemata of the butterfly schema.

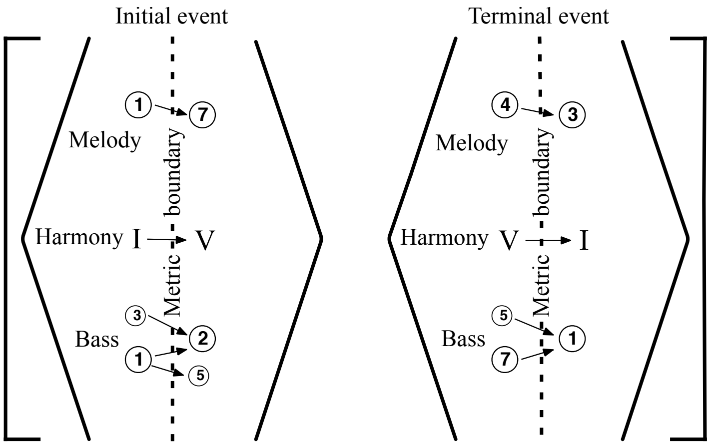


Figure 4.11 The 1–7… 4–3 schema (Gjerdingen, 1988, p. 64).

There are subtle structural modifications made to the 1–7…4–3 schema (first expounded in Gjerdingen (1988)) with its renaming to the 'Meyer' (see Figure 4.12) in Gjerdingen (2007). Indeed, owing to these differences, it is questionable if they are the same schema. The prototypical harmony of the Meyer is changed from I–V…V–I in the 1–7…4–3 schema to I–vii^{6}…V^{6}–I in the Meyer. However, in Gjerdingen (2007) Roman numerals are not used to label chords and there is a more nuanced illustration of the interaction between metre and events. Gjerdingen (2007, p. 459) points out that a significant feature of the Meyer is that its four events are 'presented in pairs at comparable locations in the meter (e.g., across a bar line, or at mid-bar, with one, two, or four measures between the pairs)'. This subtly contrasts with the description of the 1–7…4–3 schema, where events are generally appended to conformant sub-phrases (Gjerdingen, 1988, p. 88).

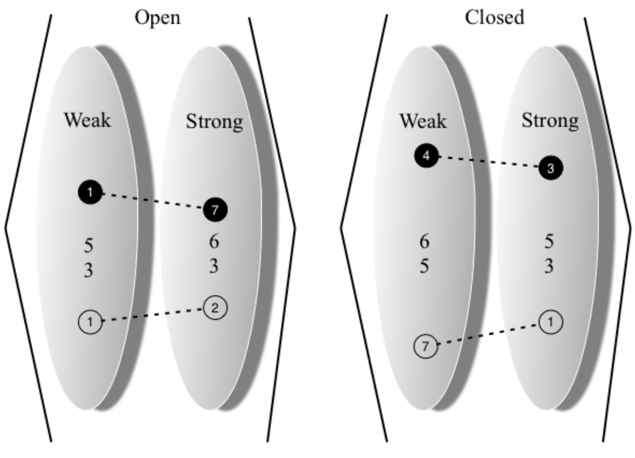


Figure 4.12 The Meyer Schema (Gjerdingen, 2007, p. 459).

More remarkable is that the depiction of the Meyer schema, shown in Figure 4.12 (Gjerdingen, 2007, p. 459), effectively places schema events in the opposite metrical positions to the grammatical depiction of the Meyer as construed by Lerdahl (2001, p. 238) (see Figure 2.8). In Lerdahl (2001, p. 238), the metrical structure of the Meyer is strong–weak…strong–weak, whereas Gjerdingen (2007, p. 459) defines the metrical structure as weak–strong…weak–strong. (In Gjerdingen (1988) the schema events straddle a metrical boundary, which is compatible with both Lerdahl (2001) and Gjerdingen (2007).) These contrasting interpretations of the metricality of schema events hinges on the different ontological status of voice-leading features in the respective theories. Similar to the present theory, although not made explicit, Lerdahl (2001) considers events to be embellishments of the grammatical structure. However, in Gjerdingen (1988, 2007), events are independent schematic structures, free from the grammatical structure, and so are not subsumed within the abstract chord progression. This means that in Gjerdingen (1988, 2007) events can be weak or strong depending on which point of the metrical structure they fall. However, in Lerdahl (2001) and the present theory, voice-leading events correspond with harmonic structure. From this perspective, events can be treated as Komar (1971) treats appoggiaturas and suspensions, as dissonances that ultimately supervene on harmonic and metrical structure. This means that melody and bass structures can be subsumed within the harmonic schema.

Using this grammatical framework, many of Gjerdingen's voice-leading schemata can be viewed as child schemata of the parent butterfly schema (shown in its preliminary version in Figure 4.9). The 1–7…4–3 of Gjerdingen (1988) (Figure 4.11) and the Meyer, Jupiter, Aprile, Pastorella, Sol-Fa-Mi, and Do–Re…Re–Mi schemata of Gjerdingen (2007) (Figures 4.12 and 4.13) are generalisable as butterfly schemata.

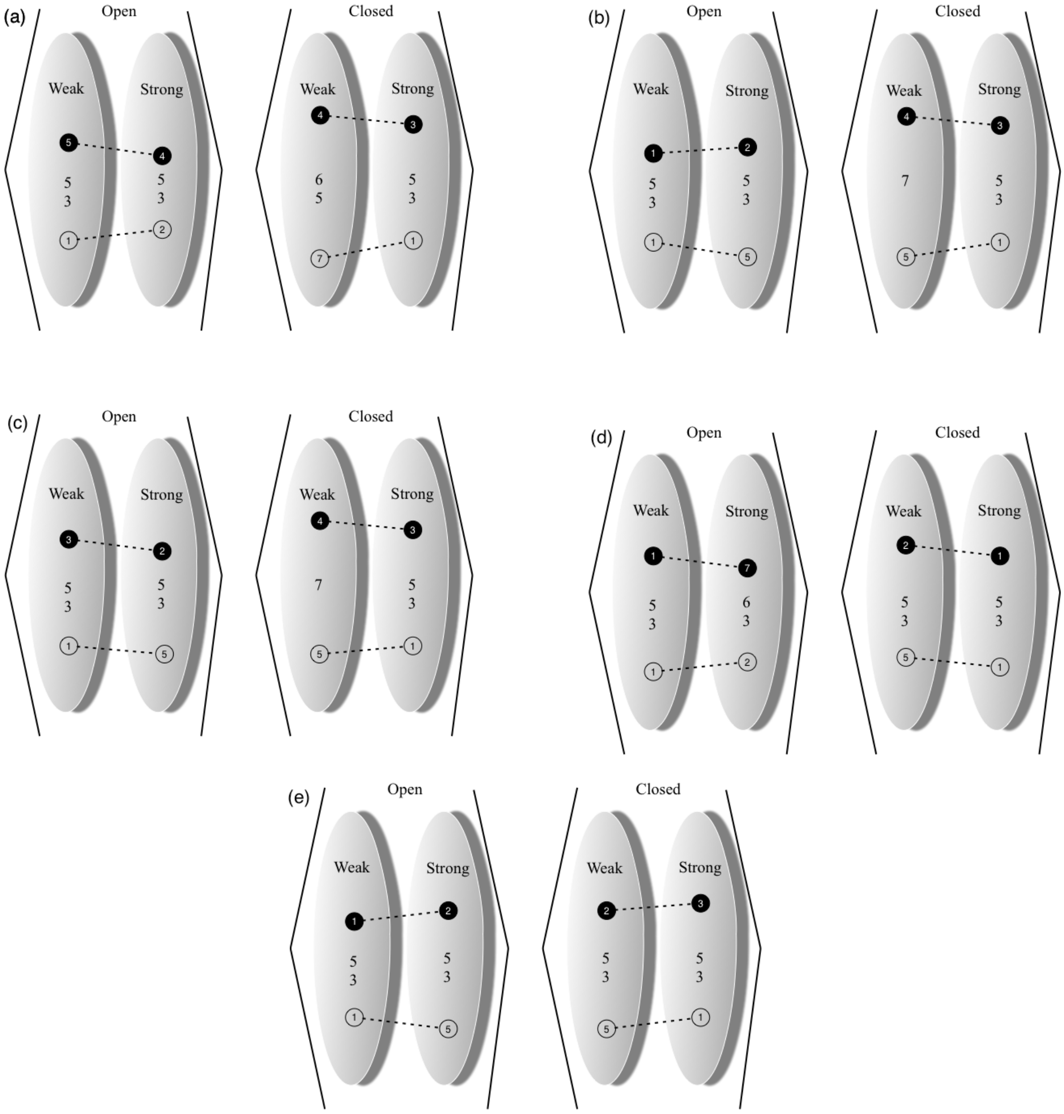


Figure 4.13 The Sol-Fa-Mi (a), the Jupiter (b), the Pastorella (c), the Aprile (d), and the Do–Re…Re–Mi (e) (Gjerdingen, 2007, pp. 86–88, pp. 111–128).

As discussed, an important contrast with the butterfly schema model and the voice-leading schemata of Gjerdingen (1988, 2007) is the lack of attention given to the implicit multiparametrically congruent relationships between chord progression, harmonic rhythm, textural grouping, and also metrical and tonal structure. This lack of recognition of the butterfly schema as a multiparametrically structure, underpinning voice-leading schemata, arguably has significant ramifications on the feasibility of the analytical claims in Gjerdingen (1988, 2007). As will be shown in the following sections, less congruent grammatical relationships, which are not butterfly schemata, are occasionally permitted in Gjerdingen's analyses because grammatical structure is a peripheral consideration to voice-leading features. Also, congruent butterfly schemata are sometimes not permitted because these schemata lack these prescribed voice-leading features.

A number of authors have questioned the methodological principle of prioritising voice-leading events in schemata. Temperley (2001, 2006) points out salient issues for schema theory, hinting that Gjerdingen (1988) lacks a consistent methodology for schematising and validating grammatical relationships:

> Surely the first and second chord of each half [of voice-leading schemata] must be part of the same phrase; an instance of the schema with a clear phrase boundary after the first I chord would hardly seem characteristic. In the vast majority of cases Gjerdingen considers, the 1–7 and 4–3 gestures appear to be metrically parallel, although perhaps this is not absolutely necessary. (Also, the 1–7–4–3 scale degrees must presumably belong to the same contrapuntal line, though that is implicit in the fact they are in the melody.) (Temperley, 2001, p. 338)
>
> [I]t is worth noting that a number of Gjerdingen's schemata use the progression I–V…V–I: the Meyer, the Jupiter, the Pastorella, the Aprile, and the Sol-Fa-Mi. Clearly, certain scale-degree realizations of this progression were schemata in their own right, but could we not regard the I–V… V–I pattern as a more general schema in itself, one that has a number of specific scale-degree variants? (Temperley, 2006, pp. 284-285)

The question of the ontological status of schemata is really at dispute here, since the portrayal of voice-leading schemata as concrete structures conflicts with their depiction as implicit multiparametrically congruent structures (with the I–V…V–I abstract harmonic form in the case of the 1–7…4–3 schema). Temperley (2001, pp. 284–285) might be suggesting here that Gjerdingen (1988) does not provide schematisations where features are interacting at the appropriate levels of abstraction to capture the implicit grammatical structure.

Lerdahl (1991, p. 273) notes that it is debatable whether schemata, of various types, are actual patterns in musical structure, or if they might be 'theoretical fictions' (discussed also in Section 2.2.2). In Gjerdingen (1988) schema validation is not systematic, and there is no account of the variability of the schematic structures (Cavett-Dunsby, 1990; Lester, 1990; Cohn and Dempster, 1992; Lerdahl, 2001; Temperley, 2001, 2006; Kaiser, 2007a, 2007b; Jan, 2013). Cavett-Dunsby (1990, p. 84) argues that Gjerdingen's (1988) methods for identifying schemata and their variations is obscure. An important method for validating schemata in

Gjerdingen (1988) is with respect to clarity, distinctness, vividness, and prominence of schematic events. However, many examples of the 1–7…4–3 schema in Gjerdingen (1988) do not have the defining melody structure (1–7…4–3) salient in the texture. Cavett-Dunsby (1990, p. 83) also argues that Gjerdingen is 'insensitive to voice leading and gives no account of what constitutes a melody'. One of the most stubborn obstacles impeding a Gjerdingenian analysis is the requirement that listeners perceive the 'correct' voice-leading structure (a systematic method for defining the correct analysis is not presented). Cavett-Dunsby (1990, p. 83) doubts whether it is possible for listeners to perceive voice-leading events in textures since in many of Gjerdingen's analyses events often change voices or are hidden in texture.

Although Gjerdingen (1994) presents a computational model that demonstrates how specific voices in texture might be perceived,[27] there is no reason to suggest why certain voices should be privileged over others. While schemata are defined in Gjerdingen (1988, pp. 59–67) as 'associative networks', that admit flexible analyses, the paradox is that the integrity of the voice-leading structure (as shown in Figure 4.11) is never compromised in his analyses. The voice-leading features are privileged and so must be present for validation, but the abstract implicit features of harmony, grouping, and metrical structure can tolerate a high degree of variation. Indeed, Gjerdingen (1988, p. 81) contends that although the 'presence of 1–7…4–3 dyads will be considered a necessary condition for a 1–7…4–3 schema, it is in no way a sufficient condition'. Yet there is no adequate explanation about why the melody structure should be necessary (Cavett-Dunsby, 1990, pp. 83–84), nor an explanation about why variation in the other parameters is permitted.

Incompatible with a multiparametrically congruent interpretation is the antecedent-consequent relationships in phrase structure, shown in theories that describe musical 'periods' or 'sentences' in musical structure (e.g., Caplin (1998, pp. 35–58)). Theories of phrase structure assume a necessary relationship between serial grouping structures. However, since the tendency for congruence is thought to be a cause of structure (quantified through the rule of MC), notions of relationships between phrasing are actually projections of linear causality, not necessary relationships. There are no limiting conditions on exactly what constitutes an

[27] The segmentation of the voices into temporal Gestalten, as demonstrated in Temperely (2001), also identifies salient melodic streams.

antecedent phrase and what constitutes a consequent phrase. Gjerdingen (1996, p. 27) applies the terms antecedent and consequent to the phrasing of galant schemata, asserting that there is 'grammatical entailment in the strict sense, not the vague and prevalent notion "x is nicely complemented by y"'. However, the entailment that is attributed to the schemata in Gjerdingen is similar to Caplin's (1998, pp. 49–58) loose portrayal of entailment between antecedent and consequent phrases. Entailment and causation are redundant in the butterfly schema model because melodic phrases have no prescribed serial order, but are components of the flexible and abstract multi parametrically congruent structure.

4.3.2 Noncongruent and Invalid Schemata

This subsection examines the conflict between the analytical observations of voice-leading schemata in Gjerdingen (1988) and a description of schemata in terms of their MC. It is argued that Gjerdingen's schemata should be heard in terms of their multiparametrically congruent structure because this provides a basis for cognitively combining schema features, and therefore shows how it is possible that the rule of MC causes them. That is, this examination contends that MC is implicit in musical structure, and so explains how schemata should be interpreted as congruent categories. A full quantification of the implicit UC for each feature is provided Section 5.1, while this subsection shows how these UC features interact to form multiparametrically congruent structures. Many of the schema prototypes and analyses in Gjerdingen (1988, 2007) indirectly show multiparametrically congruent relationships, but a limitation of Gjerdingen's approach is that it does not focus on thisgrammatically congruent structure.

The preliminary abstract congruent butterfly schema (Figure 4.9) is the parent schema of the voice-leading schemata of Gjerdingen, illustrated in Figures 4.11–4.13. It can be seen that the butterfly schema implies the abstract structure within which Gjerdingen's voice-leading schemata are manifest. Therefore the abstract congruent butterfly schema is here used as the yardstick by which the grammatical multiparametric congruence of the voice-leading schemata analyses can be appraised. An important consideration is whether Gjerdingen (1988) provides a framework that might usefully synthesise concrete and abstract structuring, or whether his method conflicts with an understanding of the implicitly congruent grammatical order. The analyses in Gjerdingen (1988) that diverge from the grammatical structuring of the butterfly schema violate the common multiparametrically congruent structure of the Classical grammar. As discussed, many of Gjerdingen's analyses contradict the implicit grammatical structure because they sacrifice MC in favour of incorporating voice-leading events.

Figure 4.14, while not counted statistically as a 1–7…4–3 schema in Gjerdingen (1988, p. 82), is deemed to be approximate to the 1–7…4–3 prototype. Gjerdingen (1988, p. 83) concedes that the events lack 'harmonic and melodic closure' to classify it as a 1–7…4–3 schema. However, the grammatical structure diverges widely from the parent butterfly schema and also from the 1–7…4–3 prototype, which suggests that it might be a completely different schema.

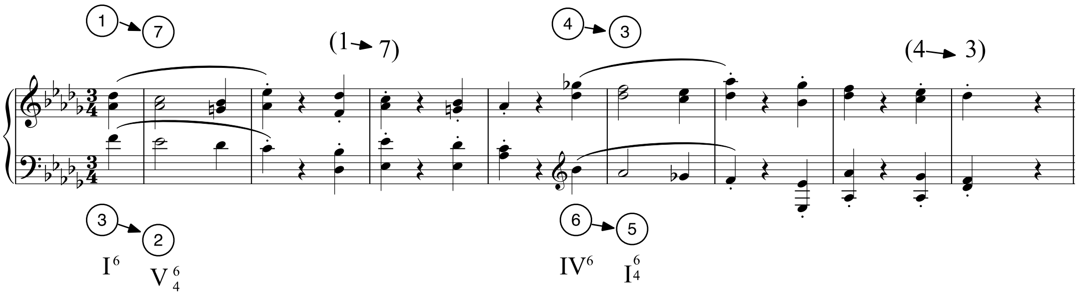


Figure 4.14 Beethoven, Piano Sonata in C-sharp minor, Op. 27, No. 2 (1801), ii, bars 1–8 (Gjerdingen, 1988, p. 82).

The grouping in bars 1–4 is parallel in rhythm and contour to the grouping in bars 5–8. The initial event in bars 1–4 follows the schematic pattern (1–7…4–3), however, in the key of A-flat major. By contrast, Gjerdingen (1988, p. 82) construes the passage as a whole in the key of D-flat major (presumably because the key signature denotes D-flat major). This presents a problem because in bars 1–4 the local tonal centre is arguably A-flat major. This is owing to the tonicisation through metrical stress of the A-flat major chords in bars 1 and 3. However, the second group is in D-flat major. Indeed, the 1–7 event is in the key of A-flat major and the 4–3 event is in D-flat major. Thus the schema events, apart from not corresponding to the norm that events are appended to the end of conformant sub-phrases (Gjerdingen, 1988, p. 88), are in different keys. This change in key conflicts with the prototype of the 1–7…4–3 schema, and the model of the butterfly schema, which must have tonic harmonic prolongation throughout. While the passage as a whole is relatively grammatically congruent, since its features are relatively uniparametrically congruent and multiparametrically congruent, the tonal structure and harmonic structure do not meet Gjerdingen's 1–7…4–3 prototype nor the type of MC required in the butterfly schema. This points to the limitation of the schema theory methodology, which privileges voice-leading events over multiparametrically congruent grammatical structure.

Figure 4.15, showing Ordonez's Symphony in C Major, iii, bar 8, is counted statistically as a 1–7…4–3 schema in Gjerdingen (1988, p. 129, p. 281). Gjerdingen uses dotted lines to show the placement of metrical boundaries. However, it is questionable whether this passage corresponds with the prototype (in Figure 4.11) because no systematic rules for categorisation are provided. More significantly, this passage contradicts the implicit grammatical structure of the parent butterfly schema. Figure 4.16 presents an analysis of the same passage noting the implicit multiparametrically congruent grammatical structure, defining the interaction between chord progression, textural grouping, and metrical structure. The schema identified (shown in the rectangular section) has a grammatical structure that is different to the congruent grammatical textural grouping and metrical structure of the butterfly schema (see Figure 4.9). The 1–7…4–3 voice-leading structure occurs on a metrical structure that is different to the preliminary model of the butterfly schema (in Figure 4.9). The 1–7…4–3 schema identified in Figure 4.16 is therefore not a butterfly schema because it does not have the particular type of congruent grammatical structure required. While the passage as a whole is grammatically congruent, this represents a different implicit structure.

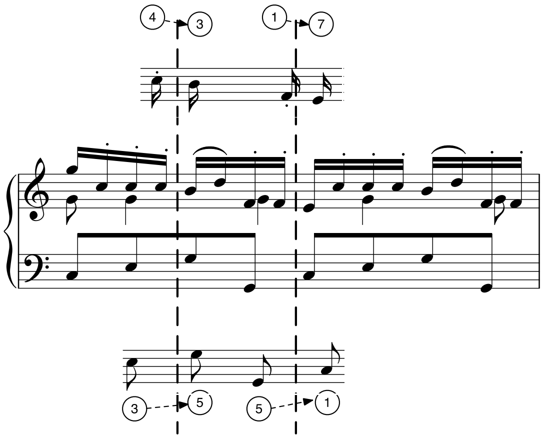


Figure 4.15 The 1–7…4–3 Schema in Ordonez, Symphony in C Major, iii (?1753), bar 8 (Gjerdingen, 1988, p. 126).

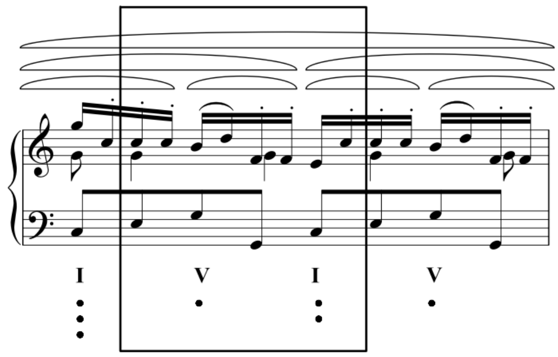


Figure 4.16 The 1–7…4–3 schema in the context of congruent grammatical structure in Ordonez, Symphony in C Major, iii (?1753), bar 8, adapted from Gjerdingen (1988, p. 126).

A further consideration is that the 1–7…4–3 schema identified in Figure 4.16 is too short to properly establish multiparametrically congruent textural grouping, chord progression, and harmonic rhythm. These features do not interact over a sufficiently long time to invoke multiparametrically congruent structure. Textural grouping, harmonic rhythm, and metrical structure must take place over a significantly long time-span to be established. At least three levels of metrical structure are required for MC between chord progression, harmonic rhythm, and textural grouping. The schema identified operates in half-beat harmonic rhythms, at the 2nd level of metrical structure, which is not extensive enough.

Figure 4.17 shows two 1–7…4–3 schemata from Veracini's Violin Sonata, Op. 1, iv. The first schema, in bars 2–4, is identified and used statistically in Gjerdingen (1988, p. 128, p. 283).

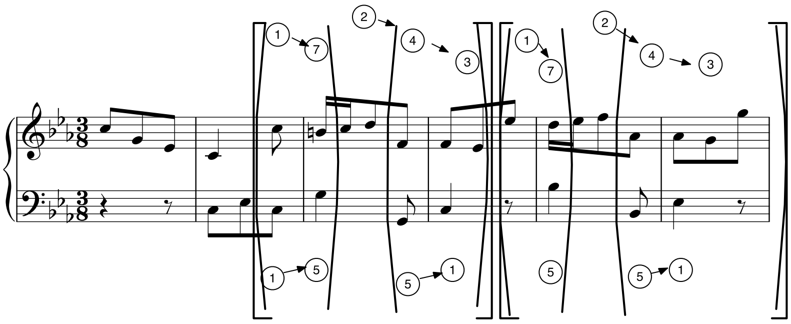


Figure 4.17 Two 1–7…4–3 schemata, Veracini, Violin Sonata, Op. 1, No. 4 (1721), iv, bars 1–6, in Gjerdingen (1988, p. 128).

The 1–7…4–3 schema in bars 2–4 of Figure 4.17 does not correspond with a multiparametrically congruent butterfly schema (see Figure 4.9). As shown in Figure 4.18, the features of the 1–7…4–3 schema in bars 2–4 do not match the multiparametrically congruent chord progression, harmonic rhythm, textural grouping, and metrical and tonal structure that is implicit in the butterfly schema, and, by extenstion, the 1–7…4–3 prototype (Figure 11). The 1–7…4–3 schema identified does not have a regular harmonic rhythm, which is required in the butterfly schema (Figure 4.9). Furthermore, the chord progression is not reinforced by metrical structure. The (weakly uniparametrically congruent) chord V is in a stronger metrical position than (the strongly uniparametrically congruent) chord I, which is a relatively noncongruent multparametric arrangement (for the butterfly schema). The melodic phrase grouping and chord progression of the first 1–7…4–3 schema identified does not correspond with the textural grouping and chord progression, which should change at the bar level of metrical structure (the level of regular functional harmonic change and at two immediately higher levels) in the butterfly schema. In sum, the placement of this proposed 1–7…4–3 schema does not match the requirements of the multiparametrically congruent butterfly schema (Figure 4.9).

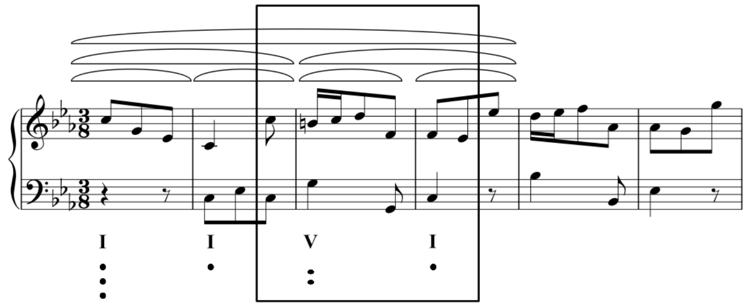


Figure 4.18 The 1–7…4–3 schema in the context of congruent grammatical structure in Veracini, Violin Sonata, Op. 1, No. 4 (1721), iv, bars 1–6, adapted from Gjerdingen (1988, p. 128).

Figure 4.19 illustrates a 1–7…4–3 schema (in bars 5–6), which is identified and used statistically in Gjerdingen (1988, p. 128). However, the passage does not correspond with the grammatically congruent parent butterfly schema (Figure 4.9). Moreover, the 1–7…4–3 schemata in this passage does not adhere to its own prototype (Figure 4.11).

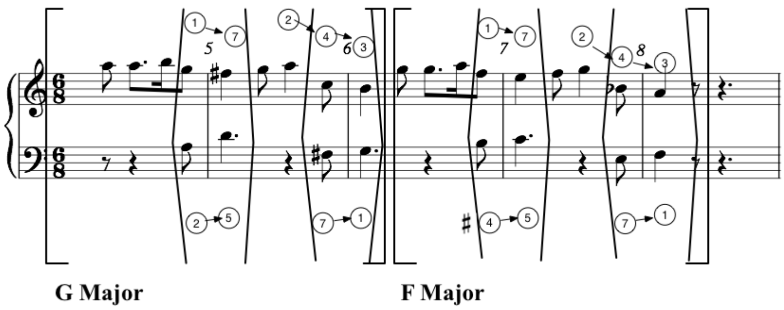


Figure 4.19 A 1–7…4–3 schema in Veracini, Violin Sonata, Op. 1, No. 2, (1721), iii, bars 4–8 (Gjerdingen, 1988, p. 128).

A multiparametrically congruent interpretation of the Veracini passage, Op. 1, No. 2, is provided in Figure 4.20. The chord progression of the 1–7…4–3 schema is ii–V…V–I, which does not correspond to the 1–7…4–3 prototype or the preliminary version of the butterfly schema. In this passage, the notated metrical structure does not reveal the actual metrical structure. The main beat is actually situated in the middle of the notated bar (bar 4), shown in Figure 4.20 with the dot structure. Bearing this in mind, the 1–7…4–3 schema is actually multiparametrically congruent with the textural grouping, chord progression, and harmonic rhythm in this passage. However, because the first event of the 1–7…4–3 schema is not underpinned by a tonic chord, the progression is not valid in terms of the 1–7…4–3 schema prototype, or for the butterfly schema. Interestingly, the melodic grouping structure does not completely correspond with the textural grouping in this passage, it is marginally out of phase, as shown in Figure 4.20. This supports the point that textural grouping is more important MC in the Classical grammar than melodic phrase grouping (examined in Section 5.1.2). That is, textural grouping is a stronger cue here for MC than melodic phrase grouping.

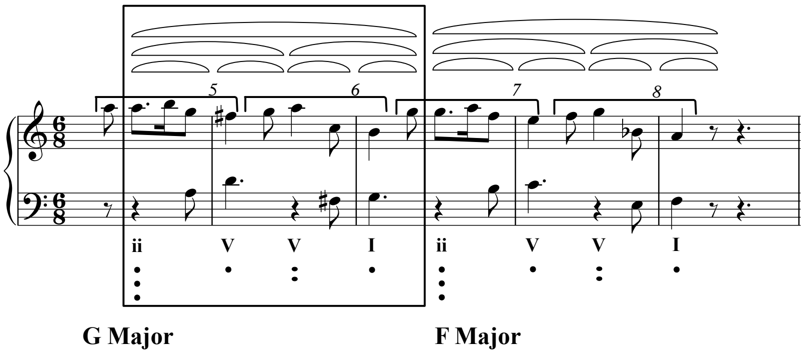


Figure 4.20 The 1–7…4–3 schema in the context of congruent grammatical structure in Veracini, Violin Sonata, Op. 1, No. 2 (1721), iii, bars 5–8, adapted from Gjerdingen (1988, p. 128).

Figure 4.21 shows a 1–7…4–3 schema in Ordonez's Symphony in C major, iii, bars 101–103, identified and used statistically in Gjerdingen (1988, p. 281). It has a number of ambiguities that preclude its validation as a multiparametrically congruent butterfly schema. Figure 4.22 illustrates the grammatically congruent structure, which does not correspond with the 1–7…4–3 schema prototype (Figure 4.11) or the preliminary version of the butterfly schema (Figure 4.9). The 1–7…4–3 voice-leading features and chord progression are noncongruent with the textural grouping and metrical structure. Also, the 1–7…4–3 schema identified constitutes chords that belong to the tonal area of G major in the first part of the schema and chords that belong to C major in the second part. This differs from the 1–7…4–3 prototype, and the model of the butterfly schema, where tonic harmony prolongation must underpin throughout.

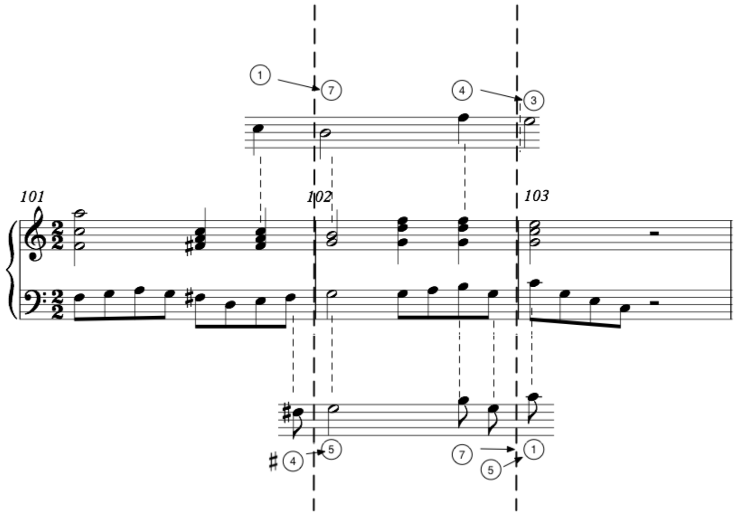


Figure 4.21 A 1–7…4–3 schema in Ordonez, Symphony in C major, Brown I:C9 (1773), iii, bars 101–103, in Gjerdingen (1988, p. 170).

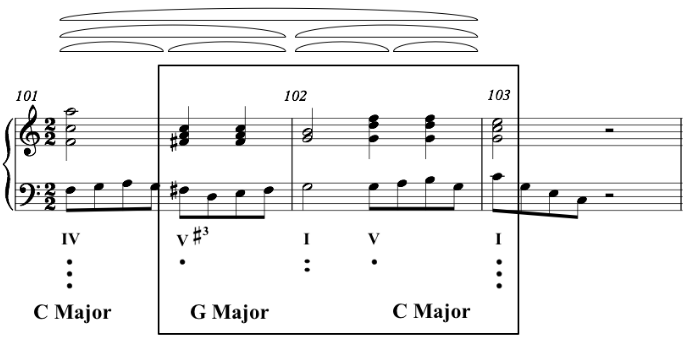


Figure 4.22 A 1–7…4–3 schema in the context of grammatical structure in Ordonez, Symphony in C major, Brown I:C9 (1773), iii, bars 101–103, adapted from Gjerdingen (1988, p. 170).

4.3.3 Noncongruent Parallel Embedded Schemata

Gjerdingen (2007, pp. 369–397) examines serial and parallel embedding of voice-leading schemata. While this practice, when considered in terms of the serial Markovian combining of schemata, is backed up by historical evidence (such as *ars combinatoria* in *partimenti* and *solfeggi*, see Sanguinetti (2012)), the parallel embedding of voice-leading schemata can be contested because it generally requires the synthesis of diverse multiparametrically congruent structures. Parallel embedding of voice-leading schemata is possible if they compliment each other grammatically. That is, if the congruent features in one schema matches the congruent features of another. In the Classical grammar this would include similar combinations of the features of chord progression, harmonic rhythm, textural grouping, as well as tonal and metrical structures. However, if the uniparametrically congruent features of these schemata conflict then they cannot be combined because the MC of one passage would not match the MC of another. A synthesis of diverse features contradicts the rule of MC, since different types of MC (which would be noncongruent with each other) cannot be generated at once. In this subsection, it is demonstrated that Gjerdingen (2007) cites instances of parallel embedding where the underlying grammatical characteristics of the respective schemata are contradictory. This suggests that neither schema might actually be present in the examples, and they show an impressionistic scattering of the features of both schemata.

The combining of similar multiparametrically congruent schemata, such as a Meyer, Aprile, Sol-Fa-Mi, Do-Re-Mi, Pastorella, or Jupiter, is feasible because they have the same congruent chord progression, harmonic rhythm, and textural grouping — these are all butterfly schemata. Only the voice-leading structures of these schemata are divergent. However, the combination of divergent grammatical schemata, as Gjerdingen (2007), means that the congruent structures must be forcibly reconciled, which is contrary to the present multiparametrically congruent conception of schemata. Figure 4.24 presents the parallel embedding (or 'piling up') of a Prinner (Figure 4.23) within a Meyer (shown in Figure 4.12) in Graun's Trio Sonata, ii, bars 1–2.

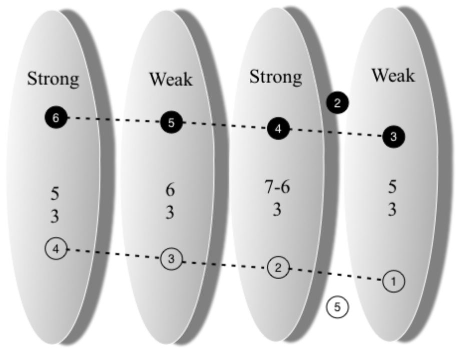


Figure 4.23 The Prinner Schema (Gjerdingen, 2007, p. 455).

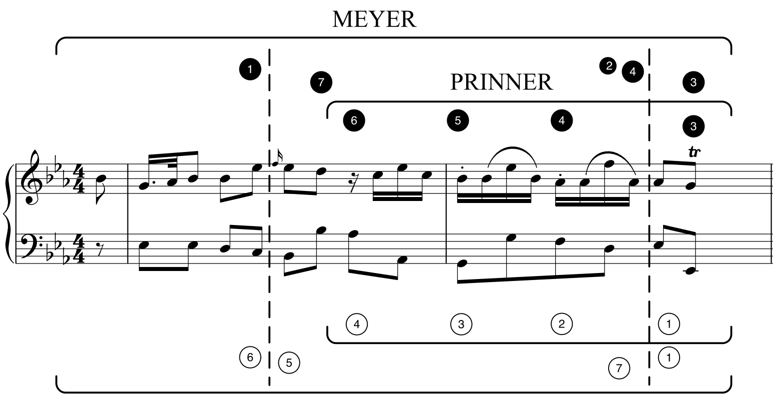


Figure 4.24 A Prinner embedded in a Meyer, Graun, Trio Sonata (*c.* 1750), ii, bars 1–2 (Gjerdingen, 2007, p. 114).

The embedding of these schemata is not possible from a perspective that values the cognitive organisation of this structure according to MC, because the multiparametrically congruent features of the Meyer, depicted as a multiparametrically congruent butterfly schema (Figure 4.9), do not match those of the Prinner (Figure 4.23). In Figure 4.25 the respective grammatical features of the original prototypes of each schema are shown in grey. It can be seen that the chord progression of the Meyer must be changed to accommodate the chord progression of the Prinner in Figure 4.25. The third chord in the sequence of a Meyer schema should be a chord V, but in the passage it is a chord I, which satisfies the Prinner but negates a grammatical Meyer. The Prinner also makes compromises to accommodate the Meyer schema in this passage, in terms of its metrical structure. The first chord of the Prinner sequence should be on a strong beat of the metrical structure (see Figure 4.23), but in this passage it is on a weak beat, which satisfies the Meyer but negates a Prinner. The actual grammatical structure of this passage is a compromise between the features of these schemata; however, it does not match either of the schemata fully. While the passage broadly complies with the rule of MC, since it is generally multiparametrically congruent — with harmonic structure close to the tonic, regular harmonic rhythm, and regular duple hierarchical textural grouping — the multiparametrically congruent structure does not correlate with the Meyer, Prinner, or butterfly schemata.

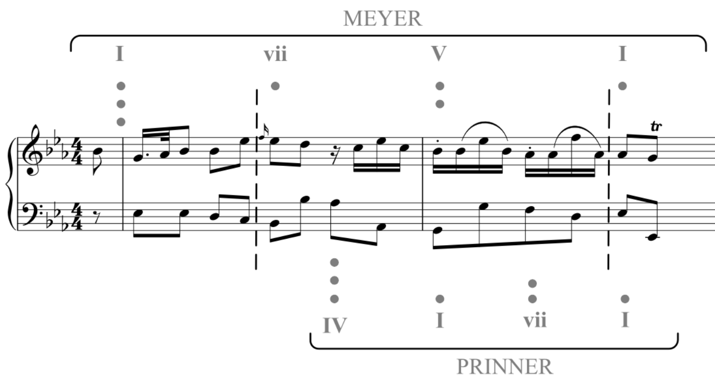


Figure 4.25 Noncongruent interaction between the Meyer and Prinner Schemata in Graun, Trio Sonata (*c.* 1750), ii, bars 1–2, adapted from Gjerdingen (2007, p. 114).

Figure 4.26 presents a similar example of contentious parallel embedding, combining a Prinner schema and a Meyer in Gluck's "Che farò senza Euridice?". Figure 4.27 shows how the multiparametrically congruent features of the prototypes of the Prinner and Meyer schemata conflict with the multiparametrically congruent grammatical structure of the passage. The third beat of bar 3 in the passage uses a chord V (Figure 4.26), but the chord sequence of the Prinner prototype (Figure 4.23) prescribes chord I. The metrical structure of the Prinner corresponds with the Meyer (assumed to be a child schema of the butterfly schema), and both follow the actual metrical structure of the passage. However, the Meyer prototype (I–vii…V–I) (Figure 4.12) requires chord V on the first beat of bar 3, but in the passage a chord IV is used, satisfying the prototype of the Prinner, but negating a Meyer. Thus, this analysis again forces two schemata into co-existence that have divergent grammatically congruent prototypes.

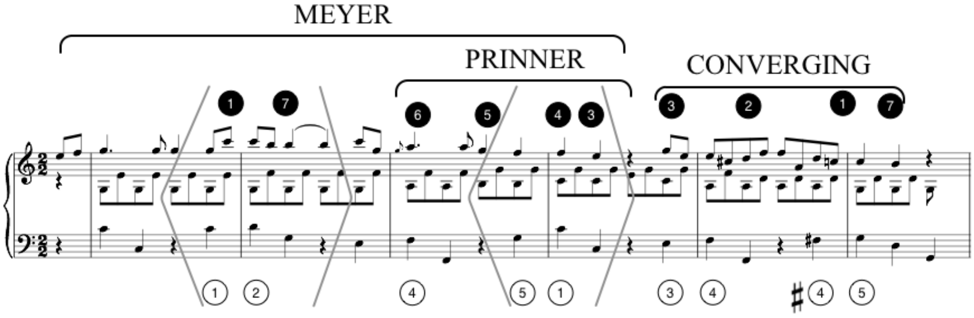


Figure 4.26 Prinner schema embedded in a Meyer, Gluck, "Che farò senza Euridice?" (1762), bars 1–6 (Gjerdingen, 2007, p. 160).

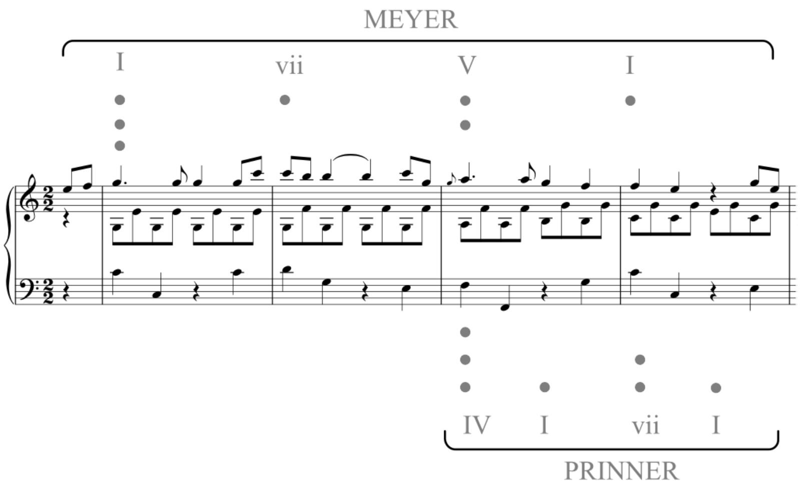


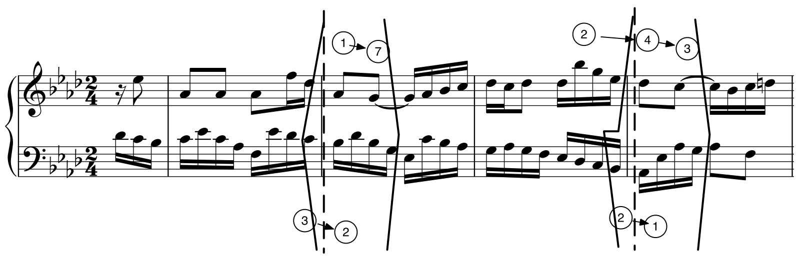
Figure 4.27 Noncongruent interaction between the Meyer and Prinner Schemata in Gluck, "Che farò senza Euridice?" (1762), bars 1–4, adapted from Gjerdingen (2007, p. 160).

4.3.4 Congruent Schemata

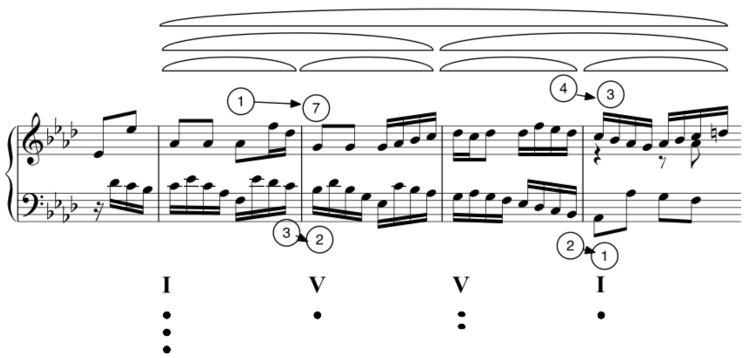
In addition to the lack of consideration of MC in Gjerdingen's analyses of 1–7…4–3 schemata, Gjerdingen (1988) fails to identify some 1–7…4–3 schemata that are multiparametrically congruent (although a consideration of MC is not central to his methodology). However, 1–7…4–3 analyses that do not observe the melody structure (1–7…4–3) are not counted in Gjerdingen's (1988) statistical analysis, even if they include the other features, such as chord progression (I–V…V–I), grouping, and metrical structure (as illustrated in Figure 4.11). The butterfly schema (Figure 4.9) uncovers the implicit multiparametrically congruent parent structure in the 1–7…4–3 prototype, and can be used to determine whether voice-leading schemata are multiparametrically congruent.

Gjerdingen (1988, p. 85) argues that the 'clarity, distinctness, vividness, or prominence' of the schema events determines whether schemata are valid. Apart from these cursory considerations a detailed framework for perceptual validation is not provided. This means that 1–7…4–3 schemata can be grammatically similar to the butterfly schema, but if they have marginally diverging melody or bass events they are not counted statistically. As discussed, chord progression, harmonic rhythm, and textural grouping form implicit multiparametrically congruent relationships in the butterfly schema, and also permit combinatorial variability in terms of the concrete elements that constitute each feature. However, melody and bass structures of voice-leading schemata are merely concrete instantiations of the implicit abstract chord structure.

Figure 4.28a shows Gjerdingen's recomposition of bars 25–28 of Bach's Fugue in F Minor from the *Well-Tempered Clavier*, Book II (1744), providing clearer and more distinct presentation of the 1–7…4–3 events than found in Bach's original. Figure 4.28b is the actual version, which is not used statistically because it is deemed to violate the 1–7…4–3 melody structure in terms of clarity and distinctness. However, the 1–7…4–3 schema in the original (Figure 4.28b) has the implicit multiparametrically congruent structure of a butterfly schema, as shown, despite its elaborative material.



a.



b.

Figure 4.28 (a) Hypothetical version of bars 25–28 of Bach's Fugue in F Minor, *The Well-Tempered Clavier*, Book II, intended to emphasize the 1–7…4–3 events; (b) actual version, bars 25–28 (Gjerdingen, 1988, p. 132), in a multiparametrically congruent butterfly schema.

The failure to count this passage as a 1–7…4–3 schema, and by implication, a butterfly schema, is probably due to the methodological focus on voice-leading events rather than the multiparametrically congruent grammatical structure. The 1–7…4–3 schema events in (Figure 4.28a) are obscured by interconnecting semiquavers (in bars 25 and 27), and are therefore presumably not recognisable in the texture. However, a clear framework for the perceptual rules for validation is not provided in Gjerdingen (1988). For example, Gjerdingen points out that,

> [i]n any 1–7…4–3 style structure, the two schema events account for only two moments. All the other music that precedes, intervenes between, or follows the two moments can effect how the schema events are perceived. For example, if a small phrase is expanded by prefacing each schema event with important melodic material, the schema events may seem to recede into the background. (Gjerdingen, 1988, p. 86)

It is not formally transparent what constitutes a case where the schema events 'recede into the background' (Gjerdingen, 1988, 86). Mozart's Piano Sonata in G Major, K. 283 (1775), i (Figure 4.29) contains a passage that is likewise not counted as a 1–7…4–3 schema in the statistical study of Gjerdingen (1988, p. 87) because of the lack of salient voice-leading events, even though it contains the required 1–7…4–3 melodic structure and multiparametrically congruent features. Indeed, despite the intervening triadic elaborations between the schema events (Gjerdingen, 1988, p. 86), the grammatical structure is a butterfly schema, as shown by the multiparametrically congruent chord progression, harmonic rhythm, and textural grouping.

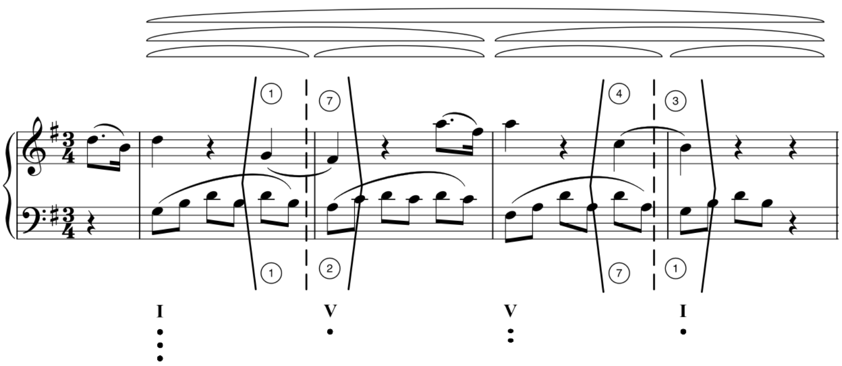


Figure 4.29 1–7…4–3 schema in Mozart, Piano Sonata in G Major, K. 283 (1775), i, bars 1–4, in Gjerdingen (1988, p. 87), in a multiparametrically congruent butterfly schema.

Figure 4.30, Mozart's Sonata in C Major, K. 309 (1777), ii, shows a further example of a 1–7…4–3 schema, which is also a butterfly schema, but is again considered invalid in

Gjerdingen (1988, p. 78) because the melody events are obscured by decorative melodic material. Also, this instance of the 1–7…4–3 schema uses a chord vii as the second chord of the progression, rather than chord V, prescribed in the 1–7…4–3 prototype (Figure 4.11). However, this chord in this position is used in the Meyer schema prototype (Figure 4.12). The covering of the final dyad of the 1–7…4–3 voice-leading feature by decorative material arguably does not diminish the multiparametrically congruent structure of the passage.

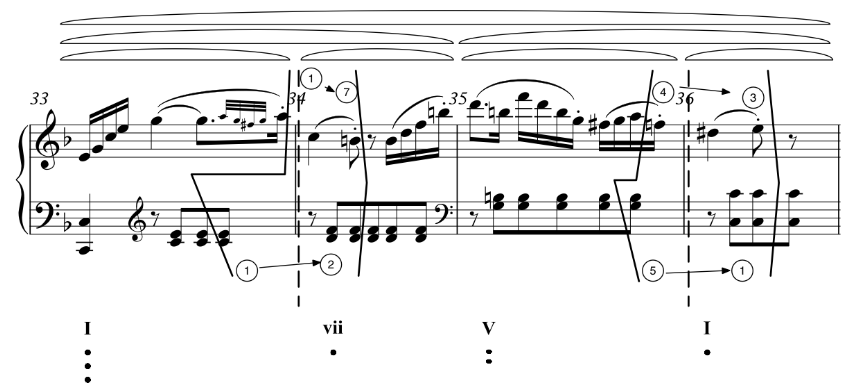


Figure 4.30 1–7…4–3 schema in Mozart, Sonata in C Major, K. 309 (1777), ii, bars 33–36 (Gjerdingen, 1988, p. 78), in a multiparametrically congruent butterfly schema.

Gjerdingen (1988, p. 84) argues that the split of the 1–7…4–3 voice-leading feature between the right and left hands of the piano in Beethoven's Piano Sonata, Op. 10, No. 3, iii (shown in Figure 4.31) hinders the perception of the opening 1–7…4–3 schema, which is grounds for it not being counted as a 1–7…4–3 schema in Gjerdingen's statistical analysis. However, the passage is a grammatically congruent butterfly schema, and arguably fits the prototype of the 1–7…4–3 schema (Figure 4.11).

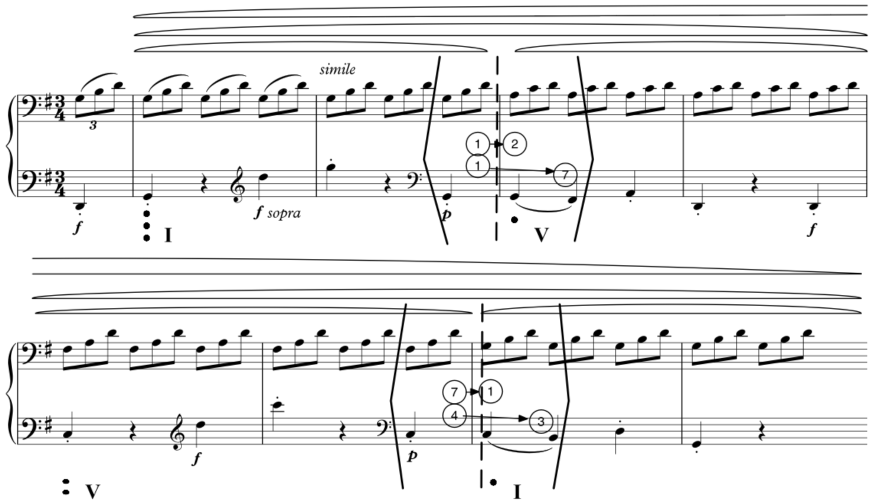


Figure 4.31 1–7…4–3 schema in Beethoven, Sonata in D Major, Op. 10, No. 3 (1797–98), iii, bars 1–8 (Gjerdingen, 1988, p. 84), in a multiparametrically congruent butterfly schema.

Figure 4.32 shows a further analysis of the 1–7…4–3 schema, embedded in a butterfly schema, in Haydn's Keyboard Sonata, Hob. XVI/23, iii, from Gjerdingen (1988, p. 70), but which is counted as invalid in his statistical analysis. The chord IV on the second beat of bar 1 is presumed not to be prototypical for the 1–7…4–3 schema because it conflicts with the chord progression of the prototype (Figure 4.11). However, since chord IV occurs on a relatively weak beat of the hypermetrical structure and resolves to a chord V in the same bar, it does not interfere with the congruent structure of the passage, since it can be construed as a 'passing chord' or even an 'appoggiatura chord'. (Appoggiatura chords are those that are positioned over a dissonant bass note and then resolved by a consonant chord with same bass note (Piston, 1941, p. 126).) Indeed, since the underlying chord structure is preserved, the interpolated chord arguably does not affect the status of the passage as a 1–7…4–3 schema and butterfly schema.

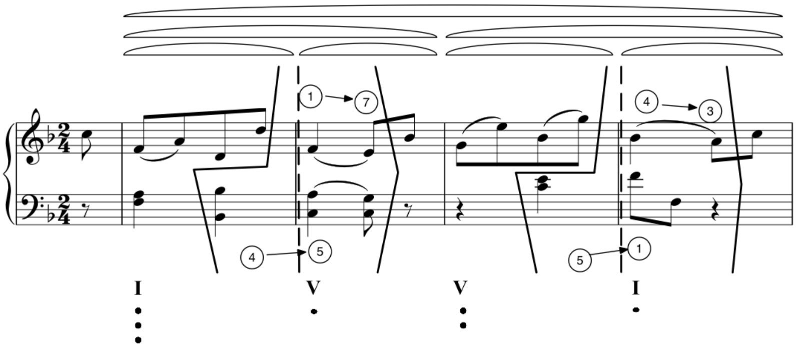


Figure 4.32 1–7…4–3 schema in Haydn, Keyboard Sonata in F Major, Hob. XVI/23 (1773), iii, bars 1–4 (Gjerdingen, 1988, p. 70), in a multiparametrically congruent butterfly schema.

4.4 Conclusions

Determining the multiparametrically congruent features of the Classical grammar and butterfly schema requires demarcating the appropriate level of abstraction where features consistently correlate. For MC, features should neither be too concrete or too abstract. This chapter demonstrates that the abstract features of chord progression, harmonic rhythm, textural grouping, and the highly abstract metrical and tonal structure, interact multiparametrically congruently in the Classical grammar and butterfly schemata. Multiparametrically congruent relationships are the implicit structure of butterfly schemata, and are also necessary for listeners to identify these schemata in real-time. The multiparametrically congruent interaction between these abstract features is implicitly understood by listeners, not learned through exposure, and thereafter applied as schemata. The butterfly schema is posited to be formed through the tendency for congruence in and between the abstract features of the Classical grammar. However, a quantitative definition of the uniparametrically congruent features and a statistical survey is necessary to provide empirical evidence for this (presented in Sections 5.1 and 5.2).

The implicit MC of the voice-leading prototypes of Gjerdingen (1988, 2007) was shown through an appraisal of his analyses. Many can be viewed as child schemata of the multiparametrically congruent parent butterfly schema. A number of 1–7…4–3 schemata identified in Gjerdingen (1988) are not butterfly schemata, and are relatively noncongruent structures, whereas other analyses fail to validate multiparametrically congruent butterfly schemata. This is because Gjerdingen’s (1988) methodology privileges melody and bass features, and does not prioritise the implicit MC. The primacy given to voice-leading events in Gjerdingen’s schema theory is in stark contrast to the anti-essentialism shown with respect to the other features of schemata, such as chord progression and metrical structure, which can be freely varied. Notwithstanding the limitations of the schema theory of Gjerdingen (1988), voice-leading events might have some significance as independent structures, but this would present an emergent level of analysis, which builds on the more fundamental congruent grammatical structure (Lerdahl, 2001, p. 248). Voice-leading schemata are better understood as scale-degree structures that are concrete instantiations of the abstract chord progression. In other grammars (such as the (pre-Classical) Galant grammar), voice-leading schemata might have more significance. However, due to the functional nature of harmony in grammars such as that of the Classical instrumental grammar, voice-leading features are generally a consequence of chords, rather than their antecedents.

Chapter 5: Modelling the Butterfly Schema

This chapter presents a model of the butterfly schemata in the Classical instrumental grammar, examining the congruent features formed by the tendency for congruence (quantified through the rule of MC) and cultural selection. To support this, the discussion includes a quantitative analysis of the uniparametrically congruent features of the Classical grammar, which are assembled into a multiparametrically congruent model. A statistical survey shows that the butterfly schema is more often formed from the grammar of the Classical period, (*c*. 1750–*c*. 1800) than the Romantic period (*c*. 1800–*c*. 1850). A hierarchical selection (HS) model of the butterfly schema and Classical grammar demonstrates that the features of the grammar and schema are culturally selected in the Classical period.

Section 5.1 examines the congruent features of the butterfly schema and Classical grammar, concentrating on the three main abstract features: chord progression, textural grouping, and harmonic-rhythm ratio.[28] The features of butterfly schemata and the Classical instrumental grammar are compared with the features of other grammars, particularly those of the Baroque and Romantic periods. These are dealt with individually in Sections 5.1.1–5.1.3, and Section 5.1.4 integrates the three uniparametrically congruent features into a unified multiparametrically congruent model.

Section 5.2 shows the distribution of butterfly schemata in the Classical period and Romantic period. A survey enables the comparison between the popularity of the butterfly schema in

[28] Chord progression, textural grouping, and harmonic-rhythm ratio are abstract features which are multiparametrically congruent with the highly abstract tonal and metrical structure.

the Classical period with that of the Romantic period in order to establish whether the tendency for congruence is the underlying cause of butterfly schemata. As discussed, it is proposed that the Classical grammar has varyingly congruent features, but tends towards congruence. If butterfly schemata are more common in the Classical period than the Romantic period then this suggests that the tendency for congruence manifest in the particular features of the Classical grammar causes butterfly schemata. Moreover, the form of butterfly schema features shows the influence of the tendency for congruence. If maximally congruent features are more common than minimally congruent features then this also suggests that they are generated by the tendency for congruence.

Section 5.3 presents a model of the butterfly schemata in the Classical grammar to show the cultural selection of features. Top-down hierarchical selection (HS) illustrates the cumulative selection of lower-level features by higher-level features, with the rule of MC constraining features in each level of the hierarchy. The particular order of dependency in top-down HS (which occurs outside the head) is the converse to the cognition of these structures in real-time listening, which is bottom-up process (that takes place inside the head) (discussed in Section 4.1.2).

5.1 A Multiparametrically Congruent Model of the Butterfly Schema

The uniparametrically congruent features of the Classical grammar and butterfly schema — the chord progression, textural grouping, and harmonic-rhythm ratio — are examined in the following subsections. The uniparametrically congruent internal structure of these features are quantified, and argued to be of a form that is particular to the Classical grammar and butterfly schema. This is determined through comparison with the congruent features of the Baroque and Romantic grammars. The uniparametrically congruent features are combined into a multiparametrically congruent model of the butterfly schema.

5.1.1 Chord Progression

The pitch-space theory of Lerdahl (2001) is incorporated to quantify the congruent chord progressions of butterfly schemata. Close proximities between chords and the tonic in diatonic space define uniparametrically congruent chord progressions of the Classical grammar and butterfly schema. Also, chord functions are used to depict the abstract chord distance categories of the butterfly schema progression. This forms a chiastic prolongational tension curve, starting on a tonic chord, departing to a relatively proximal non-tonic chord, sustaining the tension in another proximal non-tonic chord, and finally returning to the tonic chord.

5.1.1.1 Pitch-Space Theory and the Chord Functions of Schemata

Lerdahl's (2001) pitch-space theory is used to quantify the uniparametrically congruent chord progressions of the butterfly schema. The theory mathematically equates the relatedness of pitches, chords, and keys in terms of figurative spatial distance. A number of 'basic spaces' (such as diatonic space, chromatic space, etc.) are available to cognition to configure pitch-space distances. These might be generally associated with particular periods of music history, demonstrated, for example, through the increased use of chromatic space (octatonic and hexatonic) over diatonic space in the Romantic period. Regardless of the basic space in use in a grammar, listeners presumably calculate pitch distances through the most cognitively economical way, termed the 'principle of the shortest path' (Lerdahl 2001, pp. 73–77).[29] In the pitch-space theory of Lerdahl (2001), the principle of the shortest path is a cognitive strategy for listeners. However, in the butterfly schema short distances in diatonic space are preferred, forming congruent structures of the Classical grammar. In this conception of pitch-space theory, this tendency is a product of cognition rather than an explanation of cognition, quantifiable through the rule of UC.

---

[29] In terms of the tendency for congruence, the shortest path in pitch space is preferred because it is the most congruent, although this focus on congruence is not a consideration in Lerdahl (2001).

Lerdahl (2001) shows that chords can be classified into functions, which generalise the prolongational path of progressions in pitch space. As discussed in Section 2.3.7, Lerdahl (2001, pp. 193–248) shows that the voice-leading schemata of Gjerdingen (1988, 1996) can be conceived in terms of prolongational functions. Figure 5.1 shows these chord functions applied to the 1–7…4–3 schema (Lerdahl, 2001, p. 238). The T depicts a tonic, the Dep depicts a non-tonic departure from the tonic, and the Ret depicts a non-tonic return towards the tonic.

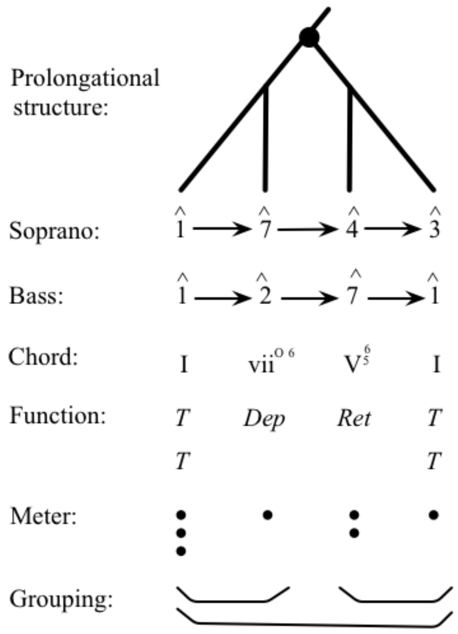


Figure 5.1 Harmonic functions of the 1–7…4–3 schema (Lerdahl, 2001, p. 238) (Repetition of Figure 2.8).

It can be observed that the 1–7…4–3 broadly uses the same chord progression, textural grouping, and harmonic-rhythm as the preliminary version of the butterfly schema (presented in Figure 4.9). The voice-leading feature of the 1–7…4–3 schema can be viewed as a child schema of the parent butterfly schema. The functions are useful because rather than representing exact distances between chords, they represent categories of tension that show how these structures are schematised in cognition (Lerdahl, 2001, p. 248). However, the chord progression in the butterfly schema model must modify these chord functions in a few important ways to correspond with the particular usage of the Classical grammar and butterfly schema. This requires a shallower tension path, representing closer distances in pitch space to the tonic, rather than the generalised functions of Lerdahl (2001), which apply to a greater variety of tonal progressions.

Butterfly schema progressions have four regions that are occupied by functional chords. Similarly to Lerdahl's model of the 1–7…4–3 schema, the first region of the butterfly schema starts on a stable tonic (T) function, which is then followed by a departure (Dep) function to higher tension, a return (Ret) function sustaining tension, and finally a stable tonic function (T–Dep…Ret–T). Following Lerdahl's model, the Dep and return Ret functions denote non-tonic chords. As pointed out, in the butterfly schema, the Dep and Ret regions must have shorter distance values from the tonic than in Lerdahl's functions, which are calculated with reference to the hypothesised patterns of usage in the Classical grammar. Short distances in pitch space are uniparametrically congruent in the butterfly schema and Classical grammar, whereas larger distances are noncongruent. Uniparametrically congruent chord progressions form consistent multiparametrically congruent relationships with other uniparametrically congruent features in the Classical grammar and butterfly schema.

It is necessary to calculate the exact distances between chords to formally define the butterfly schema functional progression. A measurement of distance between chords in the Classical grammar assumes a framework of diatonic space, which most aptly characterises its chord progressions. Diatonic space closely adheres to the 'constraints on basic spaces' of Lerdahl (2001, pp. 272–274), which govern the form of basic spaces. Indeed, diatonic space follows the constraints on basic spaces more strongly than octatonic or hexatonic space, and so could be considered a highly congruent type of basic space. It conforms to the constraints of psychoacoustics, being a cultural construction that is broadly guided by the structure of the harmonic series (Lerdahl, 2001, pp. 80–88). It is also a system that is governed by the 'hegemony of the triad' (Byros, 2009a, p. 48). The fifths level is particularly significant for determining pitch-space distance between pitch, chords, and keys (Lerdahl, 2001, p. 79), although all levels are incorporated in the comparison of pitch classes between chords. The basic diatonic space set to I/**C** is shown in Figure 5.2. This illustrates the relative strength of the octave (or root), fifth, triadic, diatonic, and chromatic levels in pitch relationships between chords.

(*a*) *octave* (*root*) *level:*	0												(0)
(*b*) *fifths level:*	0							7					(0)
(*c*) *triadic level:*	0				4			7					(0)
(*d*) *diatonic level:*	0		2		4	5		7		9		11	(0)
(*c*) *chromatic level:*	0	1	2	3	4	5	6	7	8	9	10	11	(0)

Figure 5.2 Diatonic basic space, set to I/C (C = 0, C♯ = 1, …B = 11) (Lerdahl and Krumhansl, 2007).

The diatonic chord distance rule below calculates the number of distinctive pitch classes (*k*) between chords on each level of the basic space (the octave, fifths, and triadic levels shown in Figure 5.2), the number of steps on the circle-of-fifths between the initial chord and the target chord (*x*), and the distance on the circle-of-fifths between the diatonic collection (*i*) that supports the target chord and the initial chord:

> **Diatonic chord distance rule (full version)** δ(x➔y) = *i* + *j* + k, where δ(x➔y) = the distance between chord x and chord y; *i* = the number of applications of the regional circle-of-fifths rule needed to shift the diatonic collection that supports x into the diatonic collection that supports y; *j* = the number of applications of the chordal circle-of-fifths rule needed to shift x into y; and *k* = the number of distinctive pcs in the basic space of y compared to those in the basic space of x. (Lerdahl, 2001, p. 60)

This rule measures the distance between chords (of the same key) by comparing the number of steps separating the chords on the circle of fifths (*j*) and adding their distinctive pitch classes (*k*) (*i* = 0 when chords are in the same key). What is often traditionally described as consonance or dissonance, within or between chords, is closeness or remoteness in Lerdahl's (2001) theory of pitch space. In the butterfly schema model these are interpreted as UC or non-UC.

It is necessary to consider an example of how pitch-space theory can be used to quantify chord progressions. V/C to I/C is a common progression in the Classical grammar, and often constitutes the Ret to T functional progression of a butterfly schema. The distance (δ) between I and V in the key of C is 5 (I/C to V/C) (δ(x➔y) = *i* + *j* + *k*) (0 + 1 + 4 = 5) (Lerdahl, 2007, p. 331), and the converse is the same distance. The regional distance between the chords is zero (*i* = 0) because chord V stays in the same key. The V chord moves once up the diatonic circle-of-fifths (*j* = 1). There are four resultant non-common pitch classes in each of the octave, fifth, and triadic levels in chord V compared to those of I (k = 4) (each non-common pitch class for each level is underlined in Figure 5.3) (Lerdahl, 2007, p. 332). This is a congruent progression in a butterfly schema, although clear demarcations of distance categories is required before UC can be fully understood in this grammar.

```
           7
   2       7
   2       7       11
0  2  4 5  7  9    11
0  2  4 5  7  9    11
0 1 2 3 4 5 6 7 8 9 10 11
δ (I/C→  V/C) 0 + 1 + 4 = 5
```

Figure 5.3 Pitch-space distance (δ) between chord I and V in the key of C (Lerdahl, 2007, p. 332).

A further complication must be introduced for defining butterfly schema progressions. Chord progressions of the Classical grammar can be variably congruent, since congruence is not a musical essence, but a graded concept. The present modificatons to the Dep and Ret chord functions of the butterfly schema categorise the chord progressions of this grammar and schema within a limited congruent range. The modified Dep and Ret functions represent a limited group of diatonic chords. It is useful to compare distances between the tonic and other diatonic chords in diatonic space to establish the chords used in the Classical grammar. Figure 5.4 shows the distances between the tonic chord and primary and secondary chords of the 'diatonic chordal core' in Lerdahl (2001, p. 80). All chords have a distance (δ) of between 5 and 8 from the tonic.

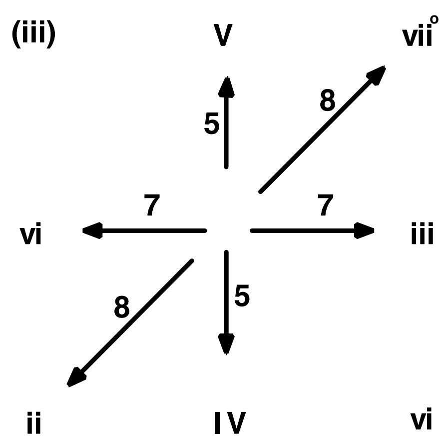


Figure 5.4 Distance values from chords of the diatonic chordal core to the tonic (Lerdahl, 2001, p. 80).

It can be observed that the distance between primary chords and the tonic (IV/iv/V ➔ I (δ = 5)) is closer than the distance between secondary chords and the tonic (vi/VI/iii/III ➔ I (δ = 7) or ii/ii°/vii° ➔ I (δ = 8)). Lerdahl (2001, pp. 77–88) notes that while these distances generally reflect practice and empirical testing, there are some issues of quantification. The theoretical distances between secondary chords and the tonic, shown in Figure 5.4, might not fully reflect empirical usage. Inconsistencies in distance values could limit the application of pitch-space theory to the butterfly schema model because chords with different pitch-space values should presumably not be functionally equivalent. For example, if chords ii, iii, and vii° are theoretically more distant from the tonic than their empirical usage might suggest (Lerdahl (2001, pp. 77–78), it is questionable whether they should be incorporated into functions that represent congruent categories.

Secondary chords are sometimes perceived as being relatively closer to the tonic in probe-tone studies (e.g., Krumhansl (1990)) than the pitch-space calculations of these distances would suggest. An explanation for this disjuncture between theory and practice is that varying interpretations of chord distances are permissible in specific contexts (Lerdahl, 2001, pp. 77–78). Chord ii often appears in first inversion, behaving like chord IV, thus having a modified distance value of $\delta = 7$, rather than $\delta = 8$ (the value of chord ii given in the chord distance chart in Figure 5.4); chord iii is often approached from upward fifth motion (rather than downward motion), increasing the distance from $\delta = 7$ to $\delta = 8$; and chord vii° often acts like chord V, lowering its chord distance value from $\delta = 8$ to $\delta = 7$ (Lerdahl, 2001, pp. 77–78). These amendments show that the distances of chords from the tonic and secondary chords can be perceived as being closer to the tonic than calculated in the chord distance rule. However, since all chords of the diatonic chordal core are relatively close to the tonic, in the distance range δ = 5–8, such differences are not detrimental for the formulation of butterfly schema functional progressions, although they might prevent difficulties should more fine-grained measurement of congruent progressions be required. An important point that may be extrapolated from this discussion is that pitch-space theory provides a useful way to provide generalised distances between chords, but measurements can vary through context and culture.

Lerdahl (2001, pp. 142–192) formulates additional components to his pitch-space theory for more nuanced distance measurements, encompassing aspects such as ‘surface dissonance’, ‘melodic attraction’, and ‘harmonic attraction’. The surface dissonance rules concern the structure of a chord, which includes the scale degree of the melody, the inversion of the chord, and additional quantification for non-chord notes (Lerdahl, 2001, p. 150). A points system is used to quantify the dissonance, which are added to or subtracted from the overall distance calculation. This arguably does not provide insight into how dissonance is cognised and interpreted, and is rather *ad hoc*. Likewise, the melodic and harmonic attraction components, while providing interesting insights, are also probably unrealistic theories about how these phenomena are understood in cognition. The attraction rules (Lerdahl, 2001, pp. 173–176) employ Newton’s inverse-square law to equate the degree of attraction between pitches and chords. Attractive relationships between pitches must surely be more complex than inverse-square relationships, involving the interaction between harmony and voice-leading.

More pointedly, the surface dissonance and melodic and harmonic attraction components in *TPS* concern more concrete levels of structure than the more abstract conceptions of functional chord relationships in the present model. Exact pitch quantifications of surface dissonance and melodic and harmonic attractions are not necessary in the butterfly schema because chords are dealt with as abstractions. It is questionable how these different approaches could be reconciled because they are different conceptual perspectives on musical structure. A further issue with the surface dissonance and melodic and harmonic attraction components is that they are not as empirically well corroborated as the diatonic chord distance rule (Lerdahl, 2008, p. 192) used to define chord functions. Due to these limitations they are not incorporated into the butterfly schema model.

5.1.1.2 The Functional Chord Progression in the Butterfly Schema

Butterfly schema chord progressions are proposed to use close pitch-space distances from the tonic in diatonic basic space. Diatonic basic space strongly adheres to the universal constraints on basic spaces (Lerdahl, 2001, pp. 272–274). Therefore the diatonic chord progressions of the Classical grammar and butterfly schema also strongly follow these universal constraints on basic spaces. As discussed, the butterfly schema uses modified Dep and Ret functions to represent only distances that are proximal to the tonic. These highly constrained diatonic constructions define the uniparametrically congruent progression of the butterfly schema model.

While short distance values between the Dep and Ret and the tonic are necessary so that butterfly schema chord progressions are highly congruent, the Dep and Ret must provide sufficient departure from the tonic ($\delta = 5$) to form a minimal tension curve (that is characteristic of this schema). The butterfly schema has low tension in the T region, moderate tension in the Dep, sustained moderate tension in the Ret, and then returns to low tension in the final T. The chord functions of the Ret and Dep regions must be limited to the 'chordal core' (Lerdahl, 2001, p. 80), in Figure 5.4, which are primary and secondary triads in diatonic space (although harmonic and melodic minor scales include non-diatonic notes, but with minimal changes in pitch distance) with a distance range of $\delta = 5$–8. It is argued that chords that have a greater distance than $\delta = 8$ are noncongruent in the butterfly schema and Classical grammar (although they might be permissible in a butterfly schema progression if they are non-functional elaborations of the functional chord progression). The formal functional progression of the butterfly schema can now be defined:

(i) *Functional chords are in a single key and key area*.

(ii) *The T regions include tonic chords*, $\delta = 0$.

(iii) *The minimal hierarchical distance between functional chords in the Dep and Ret regions and the Ts is* $\delta = 5$.

(iv) *The maximal hierarchical distance between functional chords in the Dep and Ret regions and the Ts is* $\delta = 8$.

These rules (used to validate the butterfly schema) ensure that the stable (prolongationally significant) points of the schema are at the tonics (Ts) at the beginning and end of the schema, not the central regions, Dep and Ret. The Dep and Ret regions are distant enough to be schematically pertinent or recognisable, but not too distant that they are noncongruent in the Classical grammar. Tonic chords ($\delta = 0$) in the Dep and Ret regions, or chords close to the tonic in pitch space (yielding a value lower than $\delta = 5$), even on weak beats in these regions, are not normative because this would give prolongational significance to these events. This means listeners can interpret them as resolutions or prolongations of tonic chords, weakening the tension curve of the schema. Chromatic chords in the Dep and Ret regions (exceeding a value of $\delta = 8$), are too distant from the tonic so are noncongruent. This would be uncharacteristic of the Classical grammar and butterfly schema.

The Dep and Ret functional regions have a distance range of $\delta = 5–8$. Distances within this range are valid in these regions, yielding congruent butterfly schema progressions. Valid chord progressions therefore have minimally and maximally uniparametrically congruent forms. The following minimally uniparametrically congruent form is the requirement for a valid butterfly schema chord progression:

(v) *T–Dep...Ret–T*.

By contrast, maximally congruent harmonic progressions necessitate chord V in the Ret region, which has a distance value of $\delta = 5$ from the tonic. The V in this region enables maximal UC because this region, the third quartile of the schema, enables UC to be evenly spread over the progression. Also, the V in this region permits MC with the other uniparametrically congruent features since they interact at this point. However, the Dep region is less significant for creating UC and MC; it coincides with a less significant position in the metrical structure and harmonic rhythm, on a hypermetrically weak beat. A chord V in the Dep region cannot engender maximal MC, and so it is not significant whether chords in this region are close or distant from the tonic, providing they are greater than $\delta = 5$. Thus the maximally uniparametrically congruent version uses a Dep function, but has a chord V in the third quartile:

(vi) *T–Dep…V–T*.

Arbitrary congruent chords can occupy both the Ret and the Dep regions for minimally congruent progressions, providing they are in the stipulated range (δ = 5–8). Broadly, arbitrary non-tonic chords have a similar functional role to each other. For example, chords ii, vii^{o}, and IV are common alternatives to chord V. However, as discussed, the significance of chord distance from the tonic is not as acute in the Dep region as in the Ret region, due to its position and interaction with other congruent features; the UC and MC of a chord depends on its placement in the schema.

While these variably congruent forms are useful to classify schemata, validating schemata can be difficult because chord categories are fuzzy and non-essentialist, with frequent blurring and ambiguity. For example, dissonant and interpolated notes and chords can occur within or between valid chord functions. Therefore a basic problem is how, and to what extent, non-functional chords can be permitted in the functional progression. Issues of validation can be resolved through incorporating two concepts from *GTTM*, time-span reduction and prolongational reduction. The time-span reduction selects stable pitches with a preference for rhythmic stability (tonal stability is ideally equally spaced through a piece), whereas the prolongational reduction selects rhythmic events with a preference for pitch stability (its tendency is for tonic harmony to be prolonged).

In butterfly schemata, the functions of each region are both prolongationally significant and time-span significant, but the balance between these considerations depends on the region. Generally, the functional chords are prolongationally significant in the T regions, and time-span significant in the Dep and Ret regions. In the Dep and Ret regions, functional non-tonic chords must be time-span significant because their rhythmic and metrical placement defines the function. Functional tonic chords can occupy any part of the T regions, even weak beats, because listeners cognitively prioritise the prolonged tonic. That is, the most prolongationally significant, or stable events can occur in any part of the T regions. However, stable events (δ = 0–5) must be avoided in the Dep and Ret regions because they would also be heard as prolongationally significant. This would mean the butterfly progression could not follow the characteristic shallow tension curve and become uncharacteristically stable in the middle of the schema. Distant chords, greater than $\delta = 8$, in the Dep and Ret regions would diminish the shallow tension curve of the butterfly schema because they would create deeper tension in the middle of the schema, which is unusual and noncongruent in the functional progressions of the Classical grammar. As shown in rule i, chords that involve movement to a different key or key area diminish the characteristic tension curve of the butterfly schema. These rules are incorporated into the butterfly schema model:

(vii) *Chords that are not the functional chords of Dep and Ret regions* (*including non-diatonic, passing, interpolated and appoggiatura chords*[30]) *are valid even if they are significant in the time-span reduction of those regions, occurring on the strongest beats of the regions*, *provided their resolution is on the next strongest beat in the same region as the functional chord*.

(viii) *Non-functional chords are valid in time-span significant positions in T regions provided that their resolution is to a tonic chord in those regions* (*on a weak or strong beat*).

(ix) *Tonic chords are not valid on weak or strong beats in the Dep and Ret regions*.

It would be useful to explore how these rules can be interpreted. Broadly, in the Dep and Ret regions, functional chords generally occupy the main beat of the first half of each region in duple and quadruple metres, and the main beat and the first third of the region in triple metres. Non-functional chords are permitted on the main beat of these regions if they resolve in the same region to a functional chord on the next strongest beat. This is important because non-tonic chords become significant primarily though their rhythmic and metrical placement (time-span significance). By contrast, functional tonic chords can occupy any part of the T regions, even on weak beats, because listeners prioritise the tonic in this schema and grammar in local (and perhaps global) contexts. Listeners seek tonic stability in all regions,

[30] Appoggiatura chords are chords placed over a dissonant bass note which are then resolved to form a consonance with the same bass note (following Piston (1941, p. 126)).

even when non-tonic harmony is given greater rhythmic or metrical significance. Through the incorporation of these fairly explicit rules, issues of validation and categorisation can be overcome, and noncongruent and minimally and maximally congruent functional progressions can be identified (further explored in Section 5.1.1.4).

5.1.1.3 Hierarchical and Sequential Tension in Butterfly Schemata

It is necessary to explain how pitch distance is experienced in functional progressions before it is possible to understand how functional progressions can be categorised and validated. It has been assumed in the formulations above that the measurement of distances in chord progressions, between the Dep and Ret regions and the prolonged Ts, are hierarchical. Hierarchical values are those distances between chords that have a hierarchical relationship in a prolongational context, sharing branches on a prolongational hierarchical tree (as shown in Figure 5.6). The sequential distance between two adjacent chords in a progression can also be calculated. However, listeners must generally relate chords in terms of hierarchical relationships otherwise chord and non-chord notes could not be distinguishable (See Section 1.2.1). Listening is necessarily hierarchical to make sense of musical structure (Lerdahl, 2001, pp. 152–153). Thus, when calculating pitch distances, the most stable pitches, chords, or keys must be compared across the regions of the schema. Sequential values of pitches, chords, or keys do not permit insight into how listeners hear these relationships, whereas hierarchical listening enables a retrospective and global understanding of musical structure.

Indeed, hierarchical distance calculations are necessary to model chord distances in butterfly schema progressions, and to validate those progressions. Lerdahl presents an example of a sequential analysis of distance in a I–V…V–I progression, shown in Figure 5.5. Since the Dep and Ret regions here use the same chords (chord V), the sequential distance between the Dep and the Ret is $\delta = 0$ ($\delta(x \rightarrow y) = i + j + k$) ($0 + 0 + 0$). However, this does not express the prolongational tension felt between the chord V and the hierarchically prolonged I chord in the T region. A sequential interpretation of the Ret region thus cannot account for tonic prolongation.

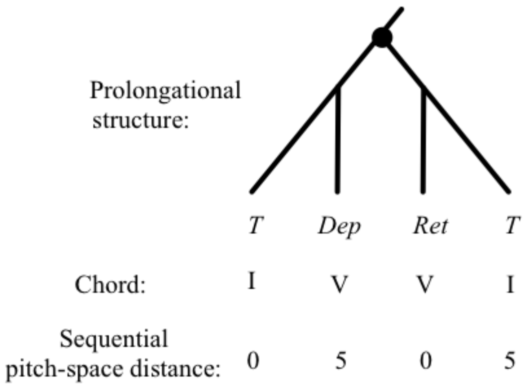


Figure 5.5 Sequential pitch-space distances of a I–V…V–I progression (based on Lerdahl (2001, p. 143)).

By contrast, a hierarchical distance analysis of the Ret region, shown in Figure 5.6, shows a distance value of $\delta = 5$ from the tonic to the dominant ($\delta(x \rightarrow y) = i + j + k$) (0 + 1 + 4). This corresponds to the experience of tension felt between the Ret region and the hierarchically prolonged tonic chord. Hierarchical listening explains how listeners perceive these prolongationally unstable chords in terms of the prolongational context. That is, hierarchical listening is necessary to understand how congruent butterfly schema progressions are defined and how they can be validated.

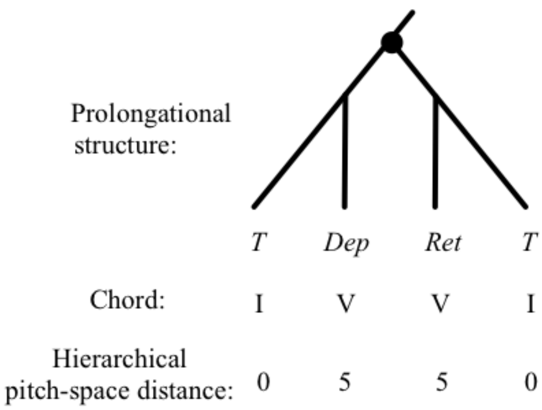


Figure 5.6 Hierarchical pitch-space distances of a I–V…V–I progression (Lerdahl, 2001, p. 146).

5.1.1.4 Issues of Validation

The diatonic chord distance rule (Lerdahl, 2001, p. 60) and the modified chord functions can now be used to explore issues of definition and validation in chord progressions of butterfly schemata. An unequivocally congruent schema is presented first, in Figure 5.7, which shows a maximally uniparametrically congruent chord progression (T–Dep…V–T) in a butterfly schema in Beethoven's Piano Sonata, Op. 2 No. 3, i, bars 1–4. This passage is also a maximally multiparametrically congruent butterfly schema, because the progression corresponds with maximally congruent textural grouping and maximally congruent harmonic rhythm. The chord progression has pitch-space distances of δ = 5 between the Dep and Ret regions and the outer T regions. That is, the chord distance from I➔V is δ = 5 (T➔Dep (δ = *i* + *j* + *k*) (δ = 0 + 1 + 4 = 5)), and this distance is the same inverted, i.e., V–I, although experienced as prolongational relaxing.

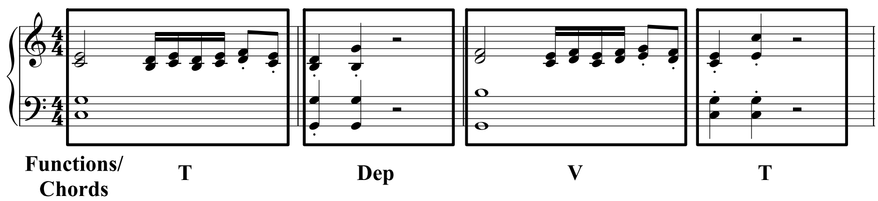


Figure 5.7 A maximally congruent chord progression in a butterfly schema in Beethoven Piano Sonata, Op. 2 No. 3 (1794–95), i, bars 1–4.

It would now be useful to consider a more distant chord progression. Figure 5.8 shows the opening of Beethoven's Sonata in C minor, *Pathétique*, Op. 13, i (1798), bars 1–2, which observes many of the features of the butterfly schema but contains a noncongruent chord progression that precludes its validation.

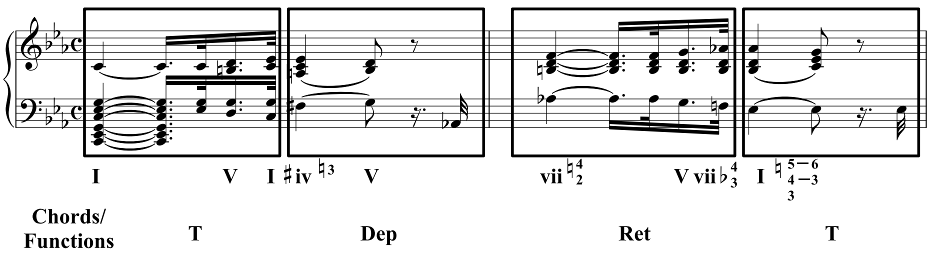


Figure 5.8 A chord progression in Beethoven Sonata No. 8 in C minor, *Pathétique*, Op. 13 (1798), i, bars 1–2.

In the Dep region, a ♯iv^{o7} chord is used that has a distance value of $\delta = 12$ from the tonic ($\delta(x \rightarrow y) = i + j + k$) $(0 + 6 + 6 = 12)$, exceeding the limit of $\delta = 8$ in this region (rule vii). It stays in the same diatonic collection ($i = 0$), moves 6 steps away from the tonic on the circle-of-fifths ($j = 6$), and has 6 non-common pitches ($k = 6$) in the levels of the spaces of the respective chords. While this ♯iv^{o7} chord is a large distance from the tonic, the chord V that follows it on the next strongest beat of the same region has a distance of $\delta = 5$ from the tonic $(0 + 1 + 4 = 5)$, providing a resolution, and making this region valid (rule vii). However, although this region is valid in this sense, the Dep region could also be viewed as a momentary movement to the dominant key area of G (vii^{o7}→I), because the leading-note of the dominant key is sharpened. Accidentals, appoggiaturas, and chromatic chords are permitted in the functional progression, providing they resolve to functional chords, but changes of key, or movement to different key areas are not (rule i). This means that the Dep region is invalid, which in the present formulation invalidates the whole passage.

The Ret region uses a diminished seventh chord (vii^{o7}), which has a distance of $\delta = 12$ $(0 + 5 + 7 = 12)$ from the tonic. The chord stays in the same diatonic collection ($i = 0$), moves 5 steps away from the tonic on the circle-of-fifths ($j = 5$), and has 7 non-common pitches ($k = 7$) in the levels of the basic spaces of the respective chords. Thus the Ret region does not satisfy the legal requirements of the schema because the vii^{o7} chord exceeds the pitch-distance limit ($\delta = 8$) (rule iv), and does not resolve to a functional chord on a relatively strong beat (rule vii).

The final T region includes a dissonant appoggiatura chord (a dissonant chord held over a note in the bass and then resolved) on the first beat, but this resolves to a tonic chord, so this region is valid (rule viii). Overall, this passage is an invalid butterfly schema because it contains momentary movement to the dominant key area G in the Dep region, and a chord that is too distant and unresolved in the Ret region. The other non-functional chords that occur on strong beats of regions are valid because they resolve to functional chords. In sum, pitch-space theory provides a way to quantify the congruent chord progressions of the Classical grammar and butterfly schemata.

5.1.1.5 Cultural Contingencies

Having presented a theoretical framework for calculating the distances in functional chord progressions of butterfly schemata, questions may be raised about whether pitch-space theory can provide a comprehensive system for predicting all patterns of chord usage in grammars and schemata. Culture is also an important constraint on the chord progressions of grammars and schemata. The chordal lexicon and patterns of usage can diverge from the universal formulations of congruent progressions, owing to the influence of culture, which is unbounded (Lerdahl, 2001, p. 381).[31] Chord types and patterns of usage are culturally transmitted. This means that although chord progressions might be constrained by diatonic space, cultural conditioning also influences them. Huron (2006, pp. 250–253) compares the distribution of chord progressions between the Western Baroque style and popular music (illustrated in Figure 5.9).

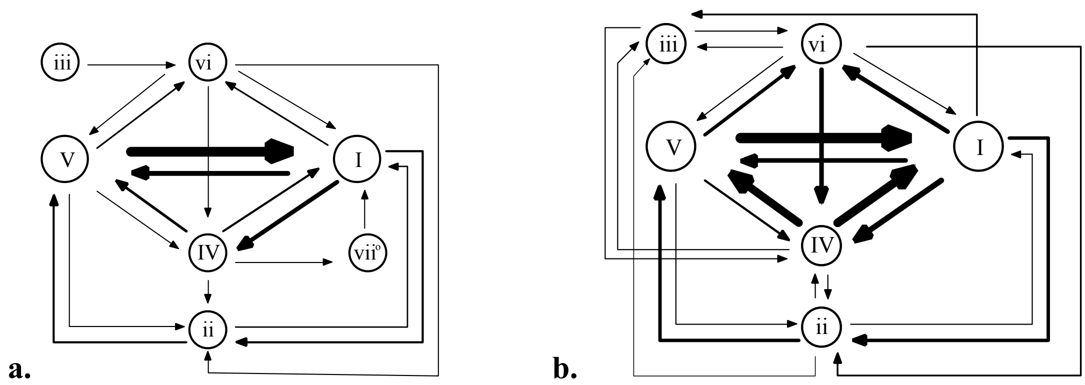


Figure 5.9 Comparison of chord progressions in (a) Baroque music; (b) Western popular songs (Huron, 2006, pp. 250–253). The thickness of the black arrows indicates the relative frequency of chord progressions in the respective styles.

[31] As discussed in Section 1.2.2, grammars are, in a sense, culturally situated even when obeying universal constraints because MC is manifest in arbitrary structures which are particular to culture.

It can be observed in Figure 5.9 that popular music has a more varied chord lexicon than baroque music (although this might seem counterintuitive).[32] The progression I–V (or V–I), which might seem to be a ubiquitous progression in Western music, is less frequent in Western popular songs, because other primary and secondary chords are also commonly used in connection to the tonic. In a similar comparative study of Baroque and reggae music (Huron, 2006, p. 203), it was found that the V–IV progression occurs twice as often as the IV–V progression in the reggae style. By contrast, the IV–V progression is much more common than the reverse, V–IV, in Baroque music. These studies in Huron (2006) show that the choice and patterning of chords is largely dependent on the culturally conditioned schemata formed in particular styles. That commonly occurring chord progressions are dependent on culture counters the notion that grammars are constrained by pitch-space distance and thus the tendency for congruence. However, the tendency for congruence does not act uniformly on cultural practices, and must be applied in particular contexts, quantified through the rule of MC (see Section 1.3). Also, there is always the potential for grammars to be designed that are largely noncongruent. Notwithstanding, the tendency for congruence is a constraint that acts in and between the features of most grammars, but in particular manifestations.

Byros (2009a, 1–9) presents further evidence for the effect of culture contradicting pitch-space universals. To this end, he explores contrasting readings of the opening of Beethoven's Symphony No. 3 in E-flat major, *Eroica*, i, Op. 55 (1804), bars 7–11, shown in Figure 5.10.

---

[32] This statistic in Huron (2006, pp. 250–253) might need explanation since Baroque harmony is often viewed as a more complex system than that of popular music. The patterns of Baroque chord usage found in Huron (2006) are presumably with reference to chord choices within local tonics, rather than within global tonics. Baroque music traverses many localised tonic regions using the cycle-of-fifths technique, however, when within these local tonic regions the chord choice is often constrained to I and V to establish the local key. Thus this seeming variety of chords in the Baroque is diminished when understood as tonic-dominant progressions of local key areas. (See Swain (2002) for a fuller discussion of surface level and abstract harmonic levels.)

Figure 5.10 Beethoven Symphony no. 3 in E-flat major, i, Op. 55 (1804), bars 1–15.

Byros (2009a) posits that there are three contrasting ways to conceive the tonality of the opening of the Eroica, in bars 5–9. Listeners can hear this section in E-flat major, G minor, or as a ‘cloud’ of indefinite tonality. Using the pitch-space calculations of chord distance in Lerdahl (2001, p. 60), and following the ‘principle of the shortest path’, these bars are universally cognised as belonging to the key of E-flat major (as calculated in Byros, 2009a). Byros (2009a) shows that a reading of the passage in the key of G minor can be made through viewing the local tonal context, and its inherent relation to the culturally situated *le–sol–fi–sol* schema. Byros’ (2009a) survey of the period shows that this schema, which broaches the tonality of the mediant minor, is found in many works of the period (supported through his statistical study). This schema might constitute ‘situated cognition’ in listeners and composers of the period. Finally, the passage can be interpreted as a ‘cloud’ of indefinite tonality, where a fixed tonal centre is not invoked in the listener. The last two types of listening are particular interpretations of the passage, which form through culturally specific listening. By contrast, the first reading is defined by the universal measure of pitch-space distance. Only listeners who are familiar with the cultural context of the passage understand the latter two types of readings. This explains how they might ‘favour the longest path’, not the ‘shortest path’ calculated through pitch-space theory (Byros, 2009a, p. 59).

This evidence of differing chord lexicons in styles and particular listening practices shows the potential for culture to override universal cognitive constraints. However, while cultural practices and ‘situated cognition’ (Byros, 2009a, p. 12) can diverge from the constraints of pitch-space theory, the incorporation of pitch-space constraints provides a universal metric by pitch structure can be quantified. If grammars are subject to the same cognitive constraints it seems doubtful that forms of situated cognition markedly differ from universal pitch-space calculations. Furthermore, relative listening practices can generally be incorporated within universal calculations of pitch space, which is a necessary basis for understanding musical structure.

Pitch-space theory calculations, based on Lerdahl (2001), can be used to measure the influence of universal constraints on grammars. Divergences from universal constraints in some grammars can provide a frame of comparison with the Classical grammar. The cultural conditioning of particular chord lexicons is not significant in the Classical grammar and butterfly schema because chord patterning conforms to the constraints that govern basic spaces (see Section 2.3.7). If the rule of MC is significant, progressions with distantly related chords, comprising noncongruent schemata, are unlikely to commonly form in the Classical grammar. Indeed, short pitch-space distances are uniparametrically congruent in the Classical grammar, and define the chord progressions of the butterfly schemata (limited to those close to the tonic in pitch space). As discussed in Sections 1.2.2 and 4.1, a further nuance in differentiating between the universal and the particular is that grammars can be considered culturally situated even if they follow universal constraints. While cognition regenerates (*wiedererzeugt* in Humboldt (1836, 93)) congruence in structure, since congruent phenomena are chosen out of possible alternatives (and congruence is manifest in a variety of arbitrary types in grammars), structures are also particular to their time and place.

5.1.2 Textural Grouping

This subsection investigates the distinct uniparametrically congruent texture of the butterfly schema and Classical grammar. 'Textural accent' is the accentuation engendered through co-incidence of accentuation in different layers of texture (Lester, 1986, p. 29). 'Textural grouping' (introduced in Chapter 4) is segmentation by textural accents; it unifies groups in texture that form a constituent hierarchy which informs metrical structure. Textural grouping is a similar to the concept of fusion and contrapuntal grouping in *GTTM*, *TPS*, and *CBMS*, but differs in its specificity, since textural grouping only occurs in and between certain levels of the metrical hierarchy. The concept of textural grouping is examined, which occurs at the level of regular functional harmonic change and at two immediately higher levels in this schema and grammar. Textural grouping forms a constituent hierarchy, where lower-level phrase groupings are contained within larger phrase groups. Textural grouping is multiparametrically congruent with other uniparametrically congruent features in the butterfly schema and the Classical grammar, following the rule of MC.[33] However, this form of textural grouping is not a uniparametrically congruent feature of some other grammars in Western tonal music, such as those of the Renaissance and Baroque periods, where textural grouping forms at different metrical levels. This variability of texture occurs because the rule of MC can be variably manifest in grammars, which is a result of cultural constraints.

5.1.2.1 Textural Profiles

Particular types of textural grouping are multiparametrically congruent with particular metrical profiles in Western music (see Section 4.1). In the Classical grammar, textural grouping is more important for creating multiparametrically congruent relationships with other features, such as metrical structure, than is melodic phrase grouping (supported in part by the theories of Cone (1968) Lester (1986) and McKee (2004)). Textural grouping at the level of regular functional harmonic change and at two immediately higher levels is proposed to be a uniparametrically congruent feature of the Classical grammar and butterfly schema which consistently interacts multiparametrically congruently with other features, following the rule of MC.

In *GTTM*, melodic phrase grouping is the main influence on metrical structure. However, while it is possible for metrical structure to be inferred this way in the Classical grammar and butterfly schema, it is arguably not as strong a cue as the combined force of the whole texture. Indeed, although metrical structure can be implicit in the note values and contours of a single line, as well as through other features of melody (Benjamin, 1984, pp. 371–372), it is not as influential as texture. Mirka (2009, pp. 133–164) bases a large part of her analytical framework on the idea that separate lines in textures can imply contrary metrical structures which have the potential for conflict. Rothstein (1995) supports the existence of concurrent

---

[33] Textural grouping cues metrical structure in real-time listening (inside the head), but the *final* cause of Classical metrical structure, its general historical and cultural cause (outside the head), is cultural transmission and selection (shown in the top-down HS model).

and contradictory metrical structures, which he refers to as ‘shadow metres’, which are metrical structures that are out of phase with the main metrical structure. Textural grouping is more significant than independent phrase grouping for creating MC with other features, and therefore defining the metrical structure. However, this is only so providing there is a dominant and coherent textural grouping in the first place, as is often the case in heterorhythmic and homorhythmic grammatical profiles. The importance of texture is also supported by McKee’s (2004) ‘law of texture’, which prioritises texture as a cue for metrical structure, showing that metrical structure is inferred primarily through combined accentuation in the accompaniment texture. While metrical structure is in part the product of the overall nested textural accents and periodicities in texture, in the Classical grammar it also relies on the input from other parametric features, such as chord progression and harmonic rhythm.

Komar (1971), Lerdahl and Jackendoff (1983), and London (2004) support the view, and argue that listeners entrain to metre, and that it is an invariant well-formed structure. Metre is thus commonly depicted as a grid, with a uniform, regular, and duple hierarchical structure. However, in practice, a metrical grid is a broad metaphor for approximately regular and hierarchical metrical structure. Moreover, metrical structure is not a concrete musical feature, but is highly abstract. Well-formed grid-like metrical structure forms at certain hierarchical levels, but is more variable at other levels (Benjamin, 1984, p. 362). Cone (1968), Lester (1986, p. 127), and Rothstein (1989, 1995) note that the correspondence between textural levels gives rise to metrical structures, of a non-essential form. Generative theories of music often do not account for the variability of textural and metrical profiles in grammars. Lester (1986, p. 252) attests that the difficulty in understanding rhythmic and metrical structure is ‘the general insistence by writers on rhythm that a phrase must project only a single accentual profile and that the task of the analyst is to find that one profile’.

Indeed, metrical structure is not completely regular or hierarchical, however, cognition tends to constrain music towards the generation of duple, regular hierarchies, following the tendency for congruence. The notion of well-formed metre and metrical entrainment in cognition cannot apply to contrapuntal grammars, since noncongruent accentuation occurs at various levels of the textural and metrical hierarchies in these systems, which cannot be unified into an overall grid. In such systems it seems unclear how metrical entrainment (as portrayed in London, 2004, p. 84) could be achieved, since entrainment to a single well-formed metrical grid is not a coherent reading of grammars that have complex, idiosyncratic, and contradictory levels of accentuation. However, cognitive entrainment *to the beat* is possible, but it is important to note that it is musical structure (outside the head) that generally gives rise to cognitive representations, not the converse.

The unification of texture proposed in *TPS* and *CBMS* is not a satisfactory solution to the problem of parsing well-formed contrapuntal metre, since this would extinguish the subtle accentuations of phrase groupings in texture that give rise to complex metrical profiles. Many textures in grammars can contain conflicting accentuation at certain levels, and so it seems that listeners are mentally able to sustain a number of conflicting structures at once, and understand the ambiguous relations, without needing to construct a well-formed grid. The theory of metrical entrainment in London (2004) and his portrayal of cyclical models cannot account for the particularity and variety of textures and metrical structures in a way that permits a sensitive reading of rhythmic structure. Even if listeners are able to entrain to the beat in a musical texture, this also can be variable, so it is questionable why entrainment should be required at any level of analysis. As noted by Cone (1968, p. 78), in the Classical grammar, metrical structure breaks down at levels higher and lower than the bar level. The metrical profile of the Classical grammar differs markedly from that of the Baroque (Cone 1968; Yeston 1976; Lester 1986; Rothstein 1989, 1995; McKee 2004). Cone (1968) proposes that the beat is the main metrical unit in Baroque music (Cone, 1968, p. 66), while the bar is the main metrical unit in Classical music (Cone, 1968, p. 72).

The idea of textural grouping broadly corresponds with the metaphor of rhythmic consonance and dissonance used by various writers (e.g., Yeston (1976), Rothstein (1989), Mirka (2009)). Rhythmic consonance is a similar concept to uniparametrically congruent textural grouping, and rhythmic dissonance describes noncongruent textural relationships. However, the present sense of textural noncongruence can be disambiguated from other uses of the term ‘rhythmic dissonance’. In Yeston (1976), dissonance occurs between mathematically indivisible melodic groups, such as the conflict between surface-level duple, triple, or quintuple groups, but the present concept of dissonance refers to the structural noncongruence inherent in and between the levels of textural and metrical structures across whole grammars. Mirka (2009) uses the notion of rhythmic dissonance to describe the noncongruent interaction between phrase groupings in texture. Likewise, this type of dissonance is not structural, but concerns textural and metrical conflict that occurs through compositional manipulations of phrase grouping in texture.

The present type of dissonance is related to the 'metrical reinterpretations' and 'successive downbeats' in Rothstein (1989), the 'overlap' in Kramer (1988), and 'metrical deletion' in Lerdahl and Jackendoff (1983). However, these terms also refer to the compositionally engineered dissonance or conflict in texture. The present meaning of rhythmic and metrical dissonance, or textural noncongruence, is intrinsic to particular grammars, occurring ubiquitously in those grammars at distinct levels of hierarchy. Textures and metrical structures manifest particular types of rhythmic consonance (UC) and dissonance (non-UC) in and between particular levels of texture. This view corresponds to the notions of culturally specific textures and metrical structures theorised in Cone (1968) and Lester (1986).

Grammars of the Renaissance and Baroque periods that use such techniques as canon and fugue have particular congruent and noncongruent levels of texture built into the hierarchical system. These grammars contradict metrical Well-Formedness Rules (MWFRs) 3 and 4 in Lerdahl and Jackendoff (1983), which prescribe regular, grid-like structuring:

> **MWFR 3** At each metrical level, strong beats are spaced either two or three beats apart.
>
> **MWFR 4** (revised) The tactus and immediately larger metrical levels must consist of beats equally spaced throughout the piece. At subtactus metrical levels, weak beats must be equally spaced between the surrounding strong beats. (Lerdahl and Jackendoff, 1983, pp. 69–72)

There is serious cause to doubt the principle of well-formed grids, and the notion that humans psychologically entrain to metres in a way that is invariant (as claimed in London (2004, p. 12, p. 84)). There are many tonal grammars that have metrical profiles that do not conform to these well-formedness rules. While metrical structure can often be grid-like in grammars, grids do not occur at every level of metrical structure. As noted above (and in Section 4.1.4), grid-like structuring is rather a constraint on metrical structuring, rather than a well-formed feature. Grammars diverge from grid-like structuring at certain levels. The textures of grammars have certain levels that are generally congruent with each other, and other levels that are generally noncongruent. The interaction of congruent levels 'closely relates to what we call musical style' (Lester, 1986, p. 127).

A numerical system is necessary to refer to each level of textural and metrical hierarchies. Figure 5.11 shows the hierarchical levels of division in metrical structure, and reveals the divergent textural groupings (and corresponding metrical structures) of Classical and Baroque grammars. A higher level (H) is expressed in terms of the immediate lower level (L) through the formula H = 2 × L. In the butterfly schema and Classical grammar, textural grouping and metrical grids are proposed to form in duple hierarchies at the level of regular functional harmonic change (usually the bar level, or 5th level) and at two immediately higher levels of metrical structure (the hypermetrical levels, usually the 6th and 7th levels). Thus textural congruence in the Classical grammar broadly follows Cone's (1968) observation that the bar level is the main metrical unit. Also, in the Baroque (and Renaissance) period, textural congruence at the beat level is predominant, which is generally the 2nd or 3rd levels of metrical structure, following the observations of Cone (1968).

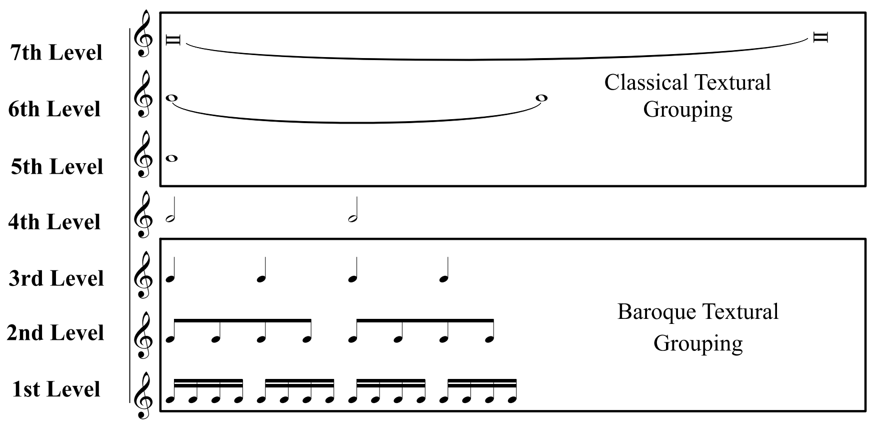


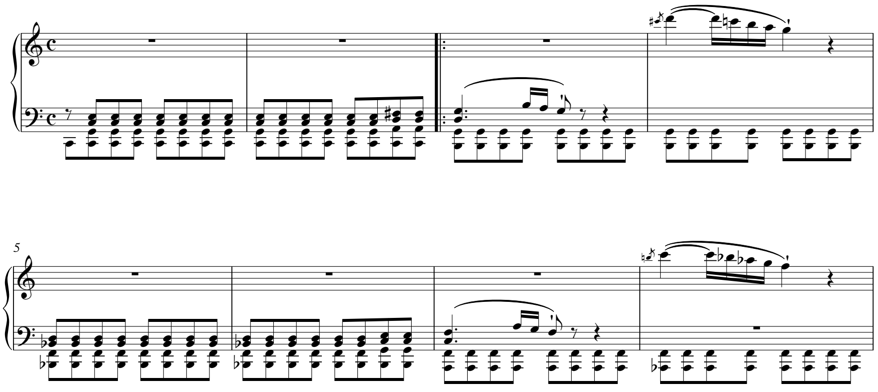
Figure 5.11 Hierarchical levels of division of metrical structure of the Classical and Baroque periods.

In contrast to Lerdahl and Jackendoff's (1983) and London's (2004) essentialist definition of metre, where a single grid is proposed to be perceptually entrained, in the present conception of metrical structure (which is supported by Hasty (1997), Lester (1986), and Cone (1968)), there can be multiple levels of beat, and also the possibility of a multiplicity of metrical structures. Metrical dissonance and even ill-formed grids can therefore be built in to metrical structures.

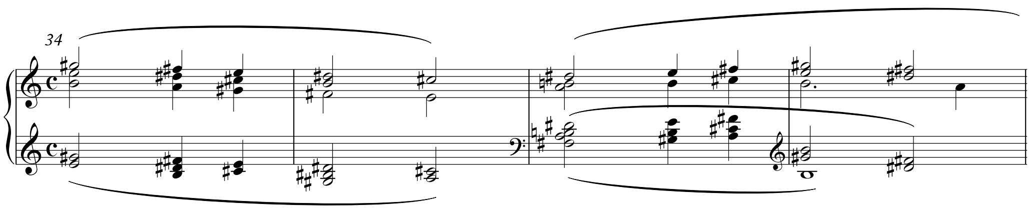
As noted in Section 4.1.4, in the Classical period, the bar level (the level of regular functional harmonic change in the present theory) is the main unit. However, the beat level might vary between sections or subsections of pieces. Cone argues that,

> […] pulse is seldom persistently obvious to our ears, and in fact the true pulse may change from one part of the movement to the next, even though the indicated meter remains the same. A movement in 4/4 may, for various themes or developments, move at a basic rate of a quarter, a half, or even a whole note […]. For it is the measure, rather than the beat, that is the fundamental unit in Classical music. The measure was important in the previous style, too, but it was to be heard as a multiplication of the primary and all-powerful beat. The beat is important in the Classical style, but it is arrived at by subdivision of the measure. That is why the beat may vary so much from one part of a movement to another: the measure is being subjected to different forms of subdivision. As a result, a Classical theme is tied more firmly to its metrical position. (Cone, 1968, p. 72)

The beat level can become metrically dissonant or noncongruent in the Classical grammar when a change of section establishes a new beat level. Indeed, the change of the beat level between themes is perhaps inherent in the sonata principle. A change of beat typically occurs between the first and second subjects of sonata form. Figure 5.12 shows the contrasting beats of the first and second subjects of Beethoven's Piano Sonata No. 21 in C major, *Waldstein*, Op. 53, (1803-1804), i. While both subjects are consistently consonant at the bar level (5th level of Figure 5.11) and hypermetrical levels (6th and 7th levels), the crotchet level (3rd level) is prominent in the first subject (Figure 5.12a) (because the harmonic rhythm articulates the crotchet level in the first bar), and the minim level (4th level) is prominent in the second subject (Figure 5.12b). Thus the main beat level in these themes is different. However, the bar level and hypermetrical levels are generally congruent in both subjects, as is typical in the Classical grammar.



a)



b)

Figure 5.12 Contrasting beat levels between the first subject, bars 1-13 (a) and second subject, bars 34-37 (b) of Beethoven Piano Sonata No. 21, Op. 53 (1803-1804), i.

5.1.2.2 Baroque Textural Profiles

The textural and metrical structures of the various grammars of the Baroque period are compared with those of the Classical instrumental grammar, to support the claim that Baroque grammars cannot generate butterfly schemata. Baroque grammars generally do not have the requisite textural grouping at the level of regular functional harmonic change and at two immediately higher levels, as does the Classical grammar. Lester (1986, pp. 127–156) uncovers an important feature of the Baroque metrical profile, which is that while the main metrical unit of Baroque grammars is the beat, multileveled structures form at higher levels. Multileveled structures are the complex and often dissonant textural interactions at higher metrical levels (Lester, 1986).[34] Lester's analysis of the multiple metrical levels of Baroque

---

[34] Lester (1986, p. 145) terms dissonant levels 'unsynchronized levels'. The term 'unsynchronized levels' is equivalent to the 'dissonant multileveled structures' or 'noncongruent multileveled structures' used in the present discussion. The distinction between 'nested versus unsynchronized levels' in Lester (1986, p. 145) is important because metrical structures manifest dissonance between particular levels.

grammars shows the different 'speeds of motion' in the texture of the chorale-with-instrumental-accompaniment form. Lester notes that,

> [a]nother type of contrapuntal texture in Bach's works features melodic voice parts (not pedals or sustained accompanimental parts) moving at dramatically different rates. In such textures, the different rates of speed of the melodic voices create the multileveled structure. Often the slowest part in such textures is a chorale or another pre-existent melody treated as a cantus firmus. (Lester, 1986, p. 131)

This feature is broadly representative of the textural profiles in Baroque grammars. Structurally interdependent higher levels of the metrical hierarchy (different speeds of motion), or dissonant, noncongruent higher levels, are common in contrapuntal music of the Baroque. Lester (1986, p. 128, p. 153) argues that the primacy of the beat in Baroque music facilitates freedom for higher multileveled pacings.

Lester (1986, p. 153) distinguishes between the contrasting profiles of Baroque and Classic-Romantic metric hierarchies, suggesting that '[t]he compositional attitudes that allow or do not allow multiple pacings to arise lie deep in the style of an era or of a composer'. An examination of the chorale-with-instrumental form in Lester (1986, pp. 131–137) focusses on the texture in Bach's Cantata No. 40, *Wachet auf, ruft uns die Stimme*, BWV 140, showing the dissonant higher-level structuring. The higher levels of texture are cued through the chorale melody, which conflict with the higher levels cued through the instrumental texture. Thus regular hypermetrical structure is not present because of the lack of correspondence between the chorale melody and the instrumental parts.

The chorale-with-instrumental-accompaniment style generally has noncongruence built in to higher levels of texture. This can be observed in a further example, in Figure 5.13, which shows the noncongruent higher metrical levels in 'O Mensch, bewein dein Sünde groß' from Bach's St Mathew Passion, BWV 244 (1727/1729), bars 17–20. Arabic numerals (the nomenclature of Rothstein (1989)) are used to show their different hierarchical metrical structures, which are generated by the independent voices. Noncongruence is created in bars 17–19, where the soprano and orchestra are half a bar out of phase with each other. In bar 20, the alto, tenor, and bass are also half a bar out of phase with the orchestra.

Figure 5.13 Noncongruent metrical layering at higher levels in 'O Mensch, bewein dein Sünde groß' from Bach *St Mathew Passion*, BWV 244 (1727/1729), bars 17–20.

Textural and metrical noncongruence occurs between various levels in many types of instrumental music in this period, but the beat level (the 2nd or 3rd level of metrical structure, in Figure 5.11) is preserved. Thus, the assertion that textural grouping generally occurs at the beat level in Baroque grammars seems probable. Also, noncongruence at levels of texture above the beat is a structural feature of Baroque grammars. However, further examples are required to establish this general trend. The opening of Bach's Fugue in C minor, BWV 847, WTC I, in Figure 5.14, shows sustained metrical noncongruence between middle levels of texture, creating a type of bi-metrical structure, while textural congruence (textural grouping) occurs at the beat level (the 3rd level of metrical structure in this case).

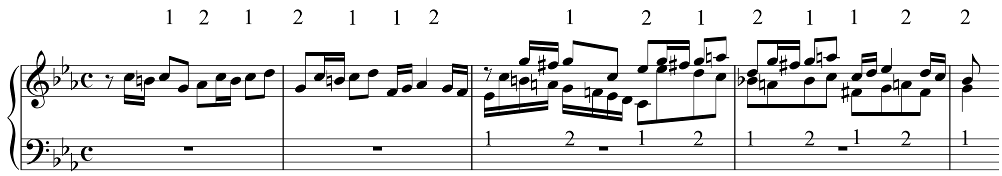


Figure 5.14 Textural noncongruence between metrical levels in the opening of Bach Fugue in C minor, BWV 847, WTC I, bars 1–5 (using the metrical nomenclature of Rothstein (1989)).

These voices create two regular, but conflicting, phrase groups. The initial middle-voice entry in bar 1 cues a metrical structure that is contrary to the notated time signature. This is achieved through accentuation of the note C″ (as opposed to the A♭′), through the use of the dactyl rhythm (see Cooper and Meyer, 1960), which makes the C″ salient by contrast to the following A♭′. Also, the primacy of the first event C″ in the middle voice gives it more salience than the following A♭′. This anchors it as a point of stability, creating a hierarchy of beats that is reinforced through repetition. The countersubject is introduced in the middle voice, beginning in bar 3, creating a hierarchy of beats corresponding with the notated metre, inferred through its contour and rhythmic groupings. This cues a metrical structure that conflicts (is noncongruent) with the fugal answer in the upper voice, at the minim level of metrical structure. The overall effect of this passage is that the accentuation at the minim level in one voice occurs simultaneously with the accentuation at the minim level in the other voice, which results in noncongruence at the respective levels. The two metrical structures are in-phase on the crotchet level (3rd level of metre), but out of phase on the minim level (4th level of metre). (The countersubject generally establishes the notated metrical structure while the subject (and answer) establishes a metrical structure that is noncongruent with this.) This type of above-the-beat textural noncongruence, and textural grouping at the beat level (congruence) is characteristic of Baroque grammars.

It should be noted that textural and metrical levels do not necessarily correlate with instrumental parts or voices, although this is often the case, as shown in *O Mensch, bewein dein Sünde groß* (Figure 5.13) and the Fugue in C minor, BWV 847 (Figure 5.14). Voices accentuate the beat level in these examples, with noncongruence at higher levels. Notwithstanding, metrical hierarchies are abstract constructions, and so voices can change levels of accentuation. While metrical structure is constructed in cognition in real-time through a bottom-up process (see Section 4.1), broad categories of abstract metrical structures (such as the higher-level dissonant structures of the Baroque period), constrained by the tendency for congruence, are transmitted in culture through top-down causation. However, the metrical structures of grammars are not necessarily limitations for composers. Regular hierarchies at particular levels are merely constraints on composers. It is possible, but generally uncommon, that specific constructions of composers can conflict with cognitive and cultural orthodoxies.

Lester (1986, p. 252), after Schenker (1935/1975), cites a further example of the noncongruent higher-level textural and metrical profile in a fugal texture (Figure 5.15), demonstrating conflicting hypermetrical structure in Bach's Fugue in C-sharp minor, BWV 849, WTC I, bars 1–11.[35] In Figure 5.15, the overlapping entries of each fugue voice creates hypermetrical dissonance, shown through Arabic numerals (following Rothstein (1989)). This passage again shows that congruence of metrical structure in Baroque grammars occurs at the beat level, but often breaks down at higher levels.

Figure 5.15 Conflicting hypermeter in Bach's Fugue in C-sharp minor, WTC I, BWV 849, after Lester (1986, p. 252), after Schenker (1935/1975), bars 1–11.

[35] This type of bi-metrical structure, with structural noncongruence in the middle and upper levels of metrical structure, is also found in various other fugues of the Baroque period, such as those of Bach's WTC, books I and II.

An important counterargument might be raised against the claim that textural congruence and noncongruence are situated in the grammars of certain periods and geographies. While contrapuntal higher-level noncongruence characterises some of the grammars of the Baroque period (and perhaps many of those of the Renaissance period), some other grammars of this period correspond more satisfactorily with the generalised conception of well-formed metre presented in generative theories. This appears to be the case with the stylised dance movements of Baroque suites, however, on a closer analysis metrical noncongruence is also intrinsic to this grammar at some higher levels. A brief exploration of the metrical profiles of Baroque dance suites reveals structures that conform to well-formed grids but are also suggestive that a well-formed grid as a blanket schema for these grammars or sub-styles is generally incoherent. In Baroque dances, each dance movement has a distinct metrical profile that only vaguely corresponds to the well-formedness rules portrayed in generative theories.

The Baroque minuet largely corresponds with the well-formed grid schema, constructed from regular textural grouping at many levels of metrical structure. However, at variance with the well-formedness of the minuet is the courante, which is a distinct Baroque French dance type, dissimilar to the Italianate corrente (Little and Jenne, 1991, p. 18). The courante has a highly specialised metrical profile with built in noncongruence at middle levels, similar to the metrical profiles of the Baroque grammars described above, such as the Bach Fugue in C minor in Figure 5.14. Little and Jenne (1991, p. 118) cite the use of 'mixed metre' in the opening of a courante by Gaspard Le Roux (1705), presented in Figure 5.16. The time signatures 6/4 and 3/2 are seemingly superimposed onto the same metrical framework, creating noncongruence between the middle levels of the textural and metrical structure. Little and Jenne note that,

> [s]everal aspects of the courante combine to create this elusive quality, chief among them the use of mixed meters …. Although the time signature is 3/2, and the first two measures are in this meter, measures 3 and 4 appear to be in 6/4, at least in the upper voice. (Measures 5 and 6 end the strain in 3/2 and 6/4 respectively, which is a common way to end a courante strain.) To add to the imbalance, measure 3 or 4, or both 3 and 4, could be read as simultaneous mixed meters, with 6/4 in the upper voice and 3/2 in the lower. (Little and Jenne, 1991, p. 118)

Figure 5.16 'Mixed metre' in a Courante by Gaspard Le Roux (1705),
Little and Jenne (1991, p. 118).

This superimposition of 6/4 and 3/2 in this dance type deviates from MWFRs 3 and 4 in the *GTTM* (cited in sub-subsection 5.1.2.1). The textural and metrical structure are noncongruent at middle levels, while maintaining congruence at the beat level, which seems to be standard in grammars of this period. Dissonant or noncongruent middle levels are common in other dances from this era, such as in the gavotte and gigue. Baroque dances in general are generally more noncongruent in the middle levels of metrical structure than dance movements of other periods, such as those of the Classical and Romantic periods, supporting the notion that this is a particular textural characteristic of this grammar.

In sum, these examples of noncongruent textural levels above the beat in Baroque profiles contradicts one of the main pillars of generative theories of music, which is the notion of well-formed metrical structure. While Lerdahl and Jackendoff (1983, p. 347) note that MWFRs 3 and 4 (stated above) are idiom-specific to tonal music, Baroque grammars — which must presumably be situated under the umbrella of tonal music — do not conform to these metrical well-formedness rules. The examples above suggest that particular ill-formed metrical grids are ubiquitous in Baroque grammars. Therefore, Baroque textures are fundamentally incapable of supporting the butterfly schema, which has a particular textural structure, as shown in the following section. While stylistic noncongrunce between levels above the beat is normative in the counterpoint of the Baroque, the Classical textural and metrical profile is well-formed at the level of regular functional harmonic change and at two immediately higher levels, necessary for the formation of the butterfly schemata. A cursory inspection of mensural structure of Renaissance grammars would show that congruence at the beat level, and noncongruence above the beat is also normative. However, these grammars require further examination, which might be undertaken in future work.

5.1.2.3 The Textural Profile of the Butterfly Schema and Classical Grammar

Textural grouping at higher levels of metrical structure is necessary for the generation of the butterfly schema. This can be formalised as a rule, which is incorporated into the butterfly schema model:

(x) *Regular, duple, and hierarchical textural grouping occurs at the level of regular functional harmonic change and at two immediately higher levels of the metrical structure*.[36]

While this type of textural grouping generally characterises the Classical grammar, the melodic phrase grouping can occasionally conflict with the main textural grouping and metrical structure. Occasional out-of-phase relationships are formed between the main metrical structure and the 'phrase rhythm' of the melody (Rothstein, 1989; McKee, 2004). This conflict provides interest and subtlety, but textural grouping is always the dominant rhythmic feature that informs metrical structure. Indeed, textural grouping is more significant than melodic phrase grouping for creating multiparametrically congruent relationships with other uniparametrically features. This contrasts with the grouping rules in *GTTM*, where melodic phrase grouping is proposed to be the primary influence on metrical structure.

This nuanced relationship between melodic phrase grouping and textural grouping requires discussion. In the Classical grammar, melodic phrase grouping can be a component of textural grouping, but can also be independent. However, while melodic phrase grouping is commonly congruent with textural grouping, it is free to conflict with textural grouping. Melodic phrase grouping is merely one voice in texture and so often it is not significant enough to override textural grouping to create MC. In Chapter 4, a reconsideration of the voice-leading schemata of Gjerdingen (and also a reframing of schematic analyses in *GTTM* and *TPS*) was suggested. When considered in terms of the framework of MC, voice-leading features can be viewed as a child schemata of the parent butterfly schema, where the melodic grouping feature is a product, not a cause, of the chord progression. Since melodic phrase

---

[36] There is occasionally noncongruence between hypermetrical levels in the metrical structure of the Classical grammar, as shown in many examples of metrical dissonances in Haydn and Mozart (see Mirka (2009)), but this is not the norm.

grouping is generally not as significant as textural grouping for MC, this provides a further argument against voice-leading schemata — because the latter are presumed to reside in melodic phrase groupings

Lester (1986, pp. 29–30) explores Classical texture in an examination of the opening of Mozart's Symphony in G Minor, K. 550 (1788), bars 1–9, shown in Figure 5.17. The passage is a maximally multiparametrically congruent butterfly schema, which is partly achieved through textural grouping at the level of regular functional harmonic change and at two immediately higher levels (rule x). The uniparametrically congruent textural grouping is shown through hemispheres, and the melodic phrase grouping is depicted by square brackets. Lester points out that,

> … [o]ne type of textural accent is caused by attacks in many or all voices of texture. These points of density are accented in relation to those points at which only one or a few voices have attacks. […] In most pieces, there is a range of textural interactions, with some relatively accented in relation to others. At the opening of Mozart's Symphony No. 40, for instance, there are three textural components: the continuous eighth-note accompaniment, the bass notes, and the melody. Attacks in all three components coincide only on downbeats in measure 2—a factor that reinforces the notated meter […]. (Lester 1986, pp. 29–30)

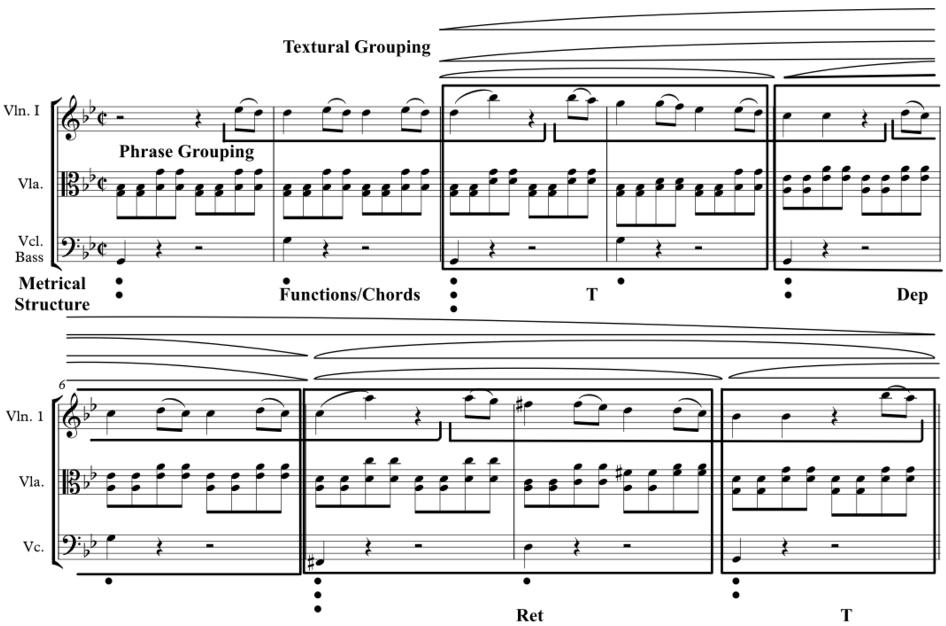


Figure 5.17 Textural grouping in a butterfly schema in Mozart Symphony in G minor, K. 550 (1788), i, bars 1–9. The textural grouping is shown with hemispheres, the melodic phrase grouping with square brackets, and the harmonic functions are separated into regions.

The melodic phrase grouping in Figure 5.17 is a bar out of phase with the textural grouping, which is a characteristic subtlety of Classical texture (Rothstein, 1995). Since the melodic phrase grouping is out of phase with textural grouping, it is therefore out of phase with the metrical structure. By contrast, textural grouping is never out of phase with metrical structure because it is the main cue for this in real-time. (However, through history and culture, metrical structure selects textural grouping, which is its top-down cause, shown in the top-down HS model in Section 5.3.)

McKee's (2004, p. 5) 'rule of texture' confirms this trend: where melodies begin before the accompaniment, the start of the accompaniment is preferably metrically stronger. With respect to this example (shown in Figure 5.17), McKee (2004, p. 4) claims that it is the accompaniment pattern that provides stable metrical organisation, following the analysis in Figure 5.17. McKee (2004, p. 4) argues that the main hypermetrical downbeat occurs after the 'extended anacrusis', in bar 3, supporting the idea that texture is multiparametrically congruent with other features in this grammar. However, McKee's interpretation, as the above reading of this passage (in Figure 5.17), differs from Lester's (1986, pp. 29–30) reading, which maintains that the main downbeat occurs on bar 2, presumably because he assumes the main metrical beat is close to the onset of the melodic phrase. Lester's interpretation is backed up by Lerdahl and Jackendoff's (1983, p. 76) MPR2 (Strong Beat Early) rule and Rothstein's (1995) 'rule of congruence'. These rules prefer correspondence between grouping and metrical structure, but they are arguably not sensitive enough to provide a coherent interpretation of the multiparametrically congruent interaction in this passage.

The analysis in Figure 5.17 (and the corresponding analysis of McKee (2004)) actually follows Rothstein's (1995, p. 173) 'rule of harmony', which posits that harmonic change corresponds with strong beats of the metrical structure. Bernstein (1976, pp. 91–105) and Temperley (2008, 2009c, pp. 134–135) argue for hypermetre beginning at bar 3, because the harmonic rhythm, which is significant for generating MC (consisting of a two-bar pattern), is congruent with the textural grouping at this point (See 5.1.3 Harmonic-Rhythm Ratios for further explanation). Moreover, the lower bass note provides further textural stability in bar 3. In sum, the present analysis, Bernstein (1976), McKee (2004), and Temperley (2009c), are arguably more coherent readings of the multiparametrically congruent interaction in this passage than Lester's (1986), although Lester (1986) provides important insights on the significance of texture, and, by extension, textural grouping.

Figure 5.18, Mozart's Symphony in A Major, KV 114 (1771), i, bars 48–53, shows a further example of the primacy of textural grouping at the level of regular functional harmonic change and at two immediately higher levels in the butterfly schema and Classical grammar. The chord progression, I–V…V–I (bars 50–52) is parallel to the melodic phrase grouping, which might give the impression of being a butterfly schema (owing to the associative network that might be suggested by its features), but these elements are noncongruent with the textural grouping (in bars 49–52) and thus noncongruent with the metrical structure. For a valid butterfly schema it is necessary that all its features are multiparametrically congruent. This structure is not a butterfly schema because it does not have the required multiparametrically congruent textural grouping. It might be described as a 'faux butterfly schema', alluding to a butterfly schema through the association of some of its features, but which actually constitute a different schema.

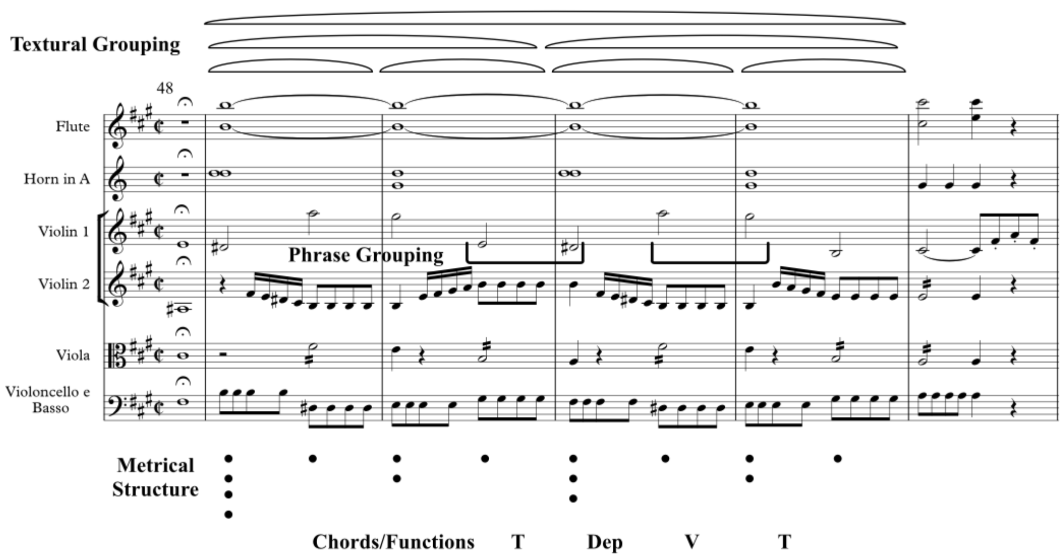
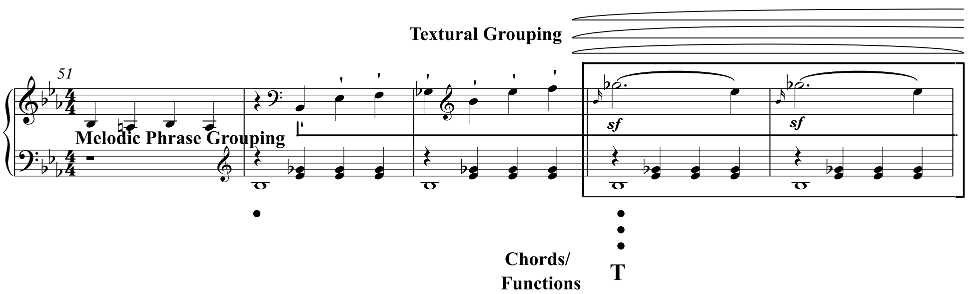


Figure 5.18 Primacy of textural grouping in Mozart's Symphony in A Major, KV 114 (1771), i, bars 48-53.

This passage is used as an example of a 1–7…4–3 schema in the statistical analysis of Gjerdingen (1988, p. 276), confirming the fundamentally different approach used in the statistical-associative formalisation of style in schema theory and the present multiparametrically congruent conception. Without considering MC there is arguably no basis for connecting the network of features in this passage (see Cavett-Dunsby (1990) and Cohn and Dempster (1992, pp. 171–172)).

While textural grouping is proposed to be more significant for producing MC than melodic phrase grouping during the Classical period, the latter is nonetheless an important component of textural grouping. Figure 5.19 shows the interdependency between melodic phrase grouping and textural grouping in a butterfly schema in Beethoven's Piano Sonata no. 8 in C minor, *Pathétique*, Op. 13 (1798), iii, bars 51–62. In this example, the bass (the low B♭ in the left hand part) and middle parts are significant in combination to create textural grouping at lower levels of metrical structure. At hypermetrical levels, however, the melody creates accentuation that cues the main hypermetrical downbeat. For example, in bar 54 the melody has an agogic accent that defines the grouping boundaries in the texture, which then cues the hypermetrical structure. A valid butterfly schema is thus generated in this passage that is largely dependent on melodic phrase grouping, but within the context of the broader structure of textural grouping. (It should be noted that harmonic rhythm also informs the hypermetrical levels in this example, dealt with in the next section.)



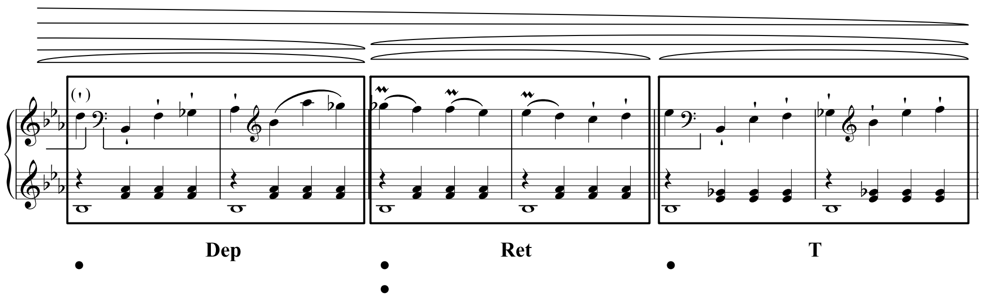


Figure 5.19 Melodic phrase grouping and textural grouping in a butterfly schema in Beethoven Piano Sonata No. 8 in C minor, Op. 13 (1798), iii, bars 51–62.

Figure 5.20 presents a butterfly schema in the opening of Beethoven's Piano Sonata Op. 2 No. 3, i (1794–95), bars 1–4, demonstrating uniparametrically congruent textural grouping at the level of regular functional harmonic change and at two immediately higher levels of metrical structure (the 5th, 6th, and 7th levels in this passage). The uniparametrically congruent textural grouping interacts multiparametrically congruently with the other uniparametrically congruent features, i.e., the chord progression and harmonic rhythm. While this type of textural grouping is a consistent component of the Classical grammar, the textural grouping can vary at high levels, to a small degree, during the course of the piece. The 6th level of metrical structure, for example, is less consistent than the 5th level in later sections of the piece (not shown in the example), where hypermeter breaks down through hypermetrical deletion or reinterpretation (such as a 2-bar hypermetrical deletion, in bars 25–26).

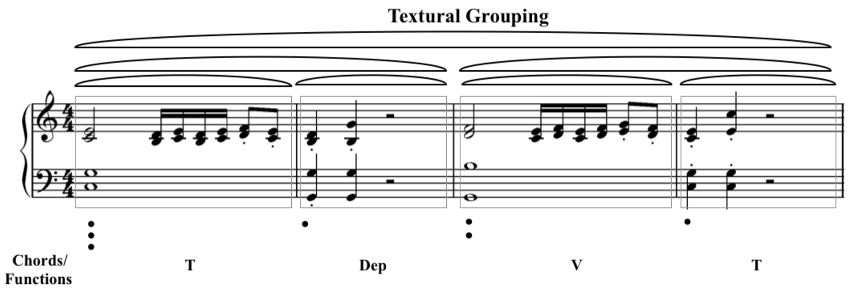


Figure 5.20 A butterfly schema with uniparametrically congruent textural grouping in Beethoven Piano Sonata Op. 2 No. 3, (1794–95), i, bars 1–4.

In summary, textural grouping is a uniparametrically congruent feature of the Classical grammar and butterfly schema. It occurs in a specific form, stated in rule x, which contrasts with the congruent textural grouping of many other grammars, such as those of the Baroque period. This uniparametrically congruent feature interacts multiparametrically congruently with the uniparametrically congruent features of the butterfly schema and Classical grammar (shown in Section 5.1.4).

5.1.3 Harmonic-Rhythm Ratio

While chord progression is often presented as an important feature of local schemata and grammars, less generally acknowledged in schema theory or generative theories is the significance of harmonic rhythm (which is the rhythm of harmonic change, explained in Piston (1941, pp. 189–203)). As also noted in Section 4.1, harmonic change is not given primary consideration in *GTTM* or *TPS*. That is, *GTTM* and *TPS* do not explore the interaction of harmonic rhythm with other parameters, aside from its general incorporation into MPR 5 Change (Mirka 2009, p. 50). However, harmonic rhythm often forms important relationships with other parametric features, as argued in Piston (1941), Berry (1976), Yeston (1976), Lester (1986), Rothstein (1995), Swain (2002), and Mirka (2009). Lester points out that,

> [h]armonic-change accentuation (harmonic rhythm) is the factor that most easily organizes pulses into a metric level. At levels where harmonic changes do not occur, or where the harmonic rhythm is ambiguous, durational and textural accents are primary factors in establishing the metric level. (Lester, 1986, p. 68)

Rothstein's (1995, p. 173) 'rule of harmony' shows the preference for changes of harmony to occur on the inception of strong beats in the metrical structure, which supports the idea that harmonic rhythm is a uniparametrically congruent feature that interacts multiparametrically congruently with other features. The use of harmonic rhythm in the present work corresponds most closely with the definition of 'phenomenal harmonic rhythm' in Swain (2002, p. 22), which is a description of the generalised functional abstraction of harmony. Harmonic rhythm is a complex notion with many subcomponents (Swain, 2002) that are too concrete to be incorporated into this presentation of the butterfly schema model (but which might be included in a finer-grained analysis of the Classical grammar in future research).

Harmonic rhythm is possibly more important in a structural sense in the Classical style than it is for many Baroque and Romantic grammars. It seems probable that functional harmonic patterns, of the type that emerge in the Classical period, cannot exist where voice-leading structuring is primary, such as in the grammars of the Renaissance and Baroque periods. In these grammars harmonic change is less often aligned with metrical structure (or mensural structure), but voice-leading considerations are more significant (Swain, 2002, p. 68). Harmonic change could therefore not create consistent multiparametrically congruent interaction with other features, which accord with the rule of MC. By contrast, in instrumental music of the Classical grammar, simple and regular patterns of harmonic rhythm are, in general, consistently aligned with textural grouping and metrical structure, following the rule of MC. For example, chords generally change once per bar in Classical grammars. Such patterning of functional harmony also exists in grammars similar to that of the Classical period, such as Western popular music, and dance music of various periods.

Swain (2002) distinguishes between music that is governed by voice-leading and that which is constrained by functional harmony:

> When a folk singer strums a guitar accompaniment, voice-leading rarely comes into play, not even unconsciously. The chords are virtually pure harmonic structures; one cannot speak of inner melodic voices, often not even in the bass. Yet the progression harmonizing the vocal melody makes sense in a way that goes beyond simply being consonant with it. The chords themselves have a coherence, an order whose justification is independent of the singer's melody and certainly independent of the voice-leading conventions associated with contrapuntal harmonic languages. This order, this purely harmonic syntax, is, of course, the hallmark of the so-called common practice period of Western art music. The elements of this syntax are harmonic functions. (Swain, 2002, p. 68)

This functional quality of harmony found in some grammars of the 'common practice period' is indicative of harmonic usage in the Classical grammar. Also, there is less independence between voices of texture. The present work is interested in this abstract notion of harmonic function, and the accentuation engendered when this interacts with rhythm and metrical structure.

Regular patterns of harmonic change are uniparametrically congruent features of the Classical grammar and butterfly schema. A novel feature introduced in the present work, the harmonic-rhythm ratio, describes the pattern of harmonic change used in the butterfly schema:

(xi) *The harmonic-rhythm ratios 1:1 and 3:1 are uniparametrically congruent features*.

Duple structuring in harmonic-rhythm ratios is simple and regular, and therefore psychologically preferred. Thus, simple and regular harmonic-rhythm ratios are uniparametrically congruent. Harmonic-rhythm ratios 1:1 and 3:1 are based on duple organisation (multiples of 2), which constitutes regular harmonic change. This permits greater hierarchical depth in the metrical structure (Lerdahl, 2001, p. 286). These ratios permit deeper hierarchical structuring than more complex ratios, such as those with underlying triple structuring (i.e., a 2:1 harmonic-rhythm ratio). Ratios 1:1 and 3:1 enable multiparametrically congruent duple relationships with textural grouping, chord progression, and tonal and metrical structure at the level of regular functional harmonic change and at two immediately higher levels. More complex ratios, such as those with underlying triple structuring, are less uniparametrically congruent, and are therefore less multiparametrically congruent.

The simplest harmonic-rhythm ratio, 1:1, is maximally uniparametrically congruent, and permits maximal MC with other uniparametrically congruent features. The harmonic change in the 1:1 ratio generally corresponds with the onsets of textural grouping and metrical structure in each region of the schema, except the third region. The third region generally prolongs the harmony of the second region, but this is accentuated in the textural grouping, so MC is still created. The 3:1 ratio is minimally uniparametrically congruent because it is less regular than a 1:1 ratio. While it permits relatively deep duple structuring and enables minimal MC with other features, it does so less congruently than the 1:1 ratio. These are formally defined and are incorporated into the model of the butterfly schema and Classical grammar:

(xii) A 1:1 harmonic-rhythm ratio is maximally uniparametrically congruent.

(xiii) A 3:1 harmonic-rhythm ratio is minimally uniparametrically congruent.

It would be useful to explore examples of harmonic-rhythm ratios in butterfly schemata. The opening of Beethoven's Piano Sonata, Op. 2 No. 3 (1794–95), i, bars 1–4, in Figure 5.21, has a maximally uniparametrically congruent harmonic-rhythm ratio (a 1:1 ratio). (Moreover, the other features of the schema are also maximally uniparametrically congruent, which means that the schema as a whole is maximally multiparametrically congruent.) By contrast, Figure 5.22 shows a butterfly schema with a minimally uniparametrically congruent harmonic-rhythm ratio (a 3:1 ratio), in Mozart's Symphony No. 39 (1788), i, bars 1–8. The other features of the schema in Figure 5.22, the chord progression and the textural grouping, are maximally uniparametrically congruent. Overall, this schema is categorised as minimally multiparametrically congruent, because this one feature is minimally uniparametrically congruent. Although congruence is a graded concept, this binary classification of the whole schema as either maximally or minimally congruent is requisite for the present quantitative analysis.

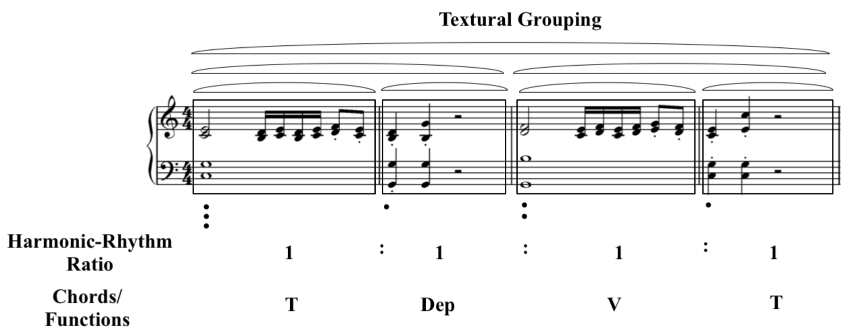


Figure 5.21 A maximally uniparametrically congruent 1:1 harmonic-rhythm ratio in Beethoven Piano Sonata Op. 2 No. 3 (1794–95), i, bars 1–4.

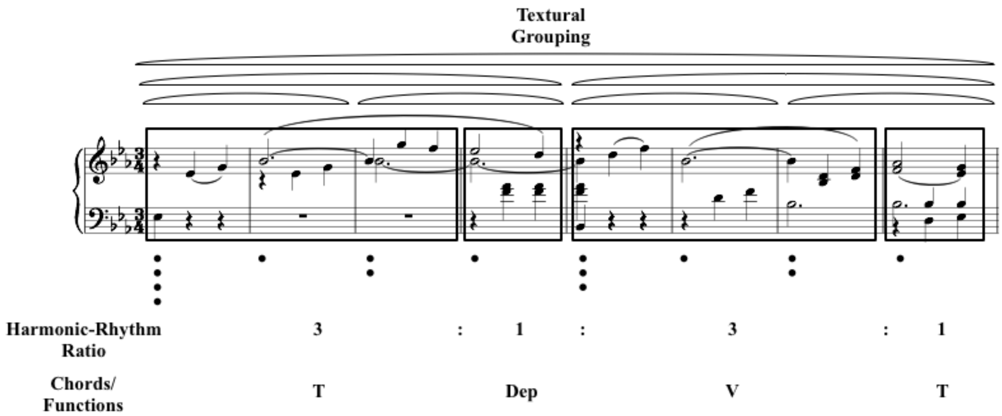


Figure 5.22 A minimally uniparametrically congruent 3:1 harmonic-rhythm ratio in Mozart Symphony No. 39 (1788), i, bars 1–8.

Figure 5.23 presents an example of a noncongruent harmonic-rhythm ratio (2:1) in the opening of Mozart's String Quartet, K. 169 (1773), ii, bars 1–6. This has the appearance of a butterfly schema on account of the maximally uniparametrically congruent chord progression and congruent textural grouping. However, a 2:1 ratio is not uniparametrically congruent in the grammar on account of its underlying triple structuring, which is complex and irregular, and not multiparametrically congruent – permitting less depth in the textural and metrical hierarchy. The 5th level of textural and metrical structure (the semibreve level) uses triple structuring, which is noncongruent with the 6th level of metrical structure (the breve level), and therefore does not form a duple hierarchy. This unbalanced harmonic rhythm engenders unbalanced textural grouping and metrical structure. In sum, simple 1:1 or 3:1 ratios, with higher-level duple structuring, are necessary for creating multiparametrically congruent hierarchies in the butterfly schema and Classical grammar.

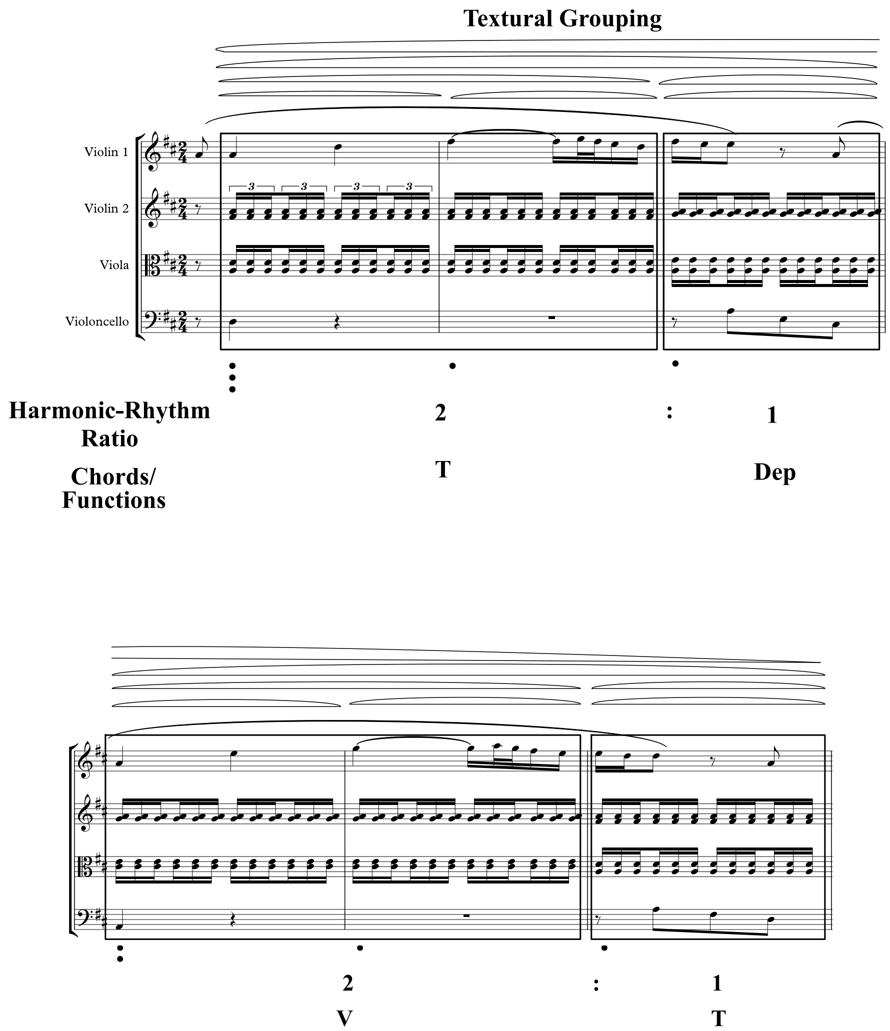


Figure 5.23 A noncongruent 2:1 harmonic-rhythm ratio in Mozart String Quartet, K. 169 (1773), ii, bars 1–6.

5.1.4 The Combined Multiparametrically Congruent Features of the Butterfly Schema

The three grammatical features shown to be uniparametrically congruent over the course of this discussion — the chord progression, textural grouping, and harmonic-rhythm ratio — can now be integrated into a multiparametrically congruent model of the butterfly schemata (assuming also multiparametrically congruent tonal and metrical structure). Figure 5.24 presents the full model:

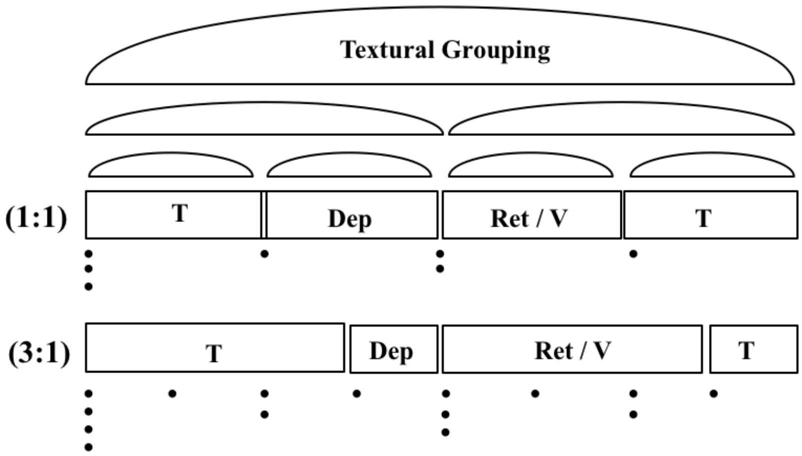


Figure 5.24 The multiparametrically congruent model of the butterfly schema.

The model includes both minimally and maximally uniparametrically congruent features because butterfly schemata are a category of variably congruent features. The chord progression and harmonic-rhythm ratios are variably congruent features, while the textural grouping is invariable. That is, the chord progression and harmonic-rhythm ratios can be minimally or maximally congruent, but textural grouping is here either congruent or noncongruent. (Textural grouping must be congruent in the butterfly schema.)

The model is sufficiently general to account for the variety of features that form in butterfly schemata as a consequence of the rule of MC. It is also particular because it describes distinct arrangements and interactions of the abstract features — chord progression, textural grouping, and harmonic-rhythm ratio — that are generally found together in the Classical grammar. Indeed, it is unlikely (but possible) that the butterfly schema exists outside Western and Western-influenced cultures. It relies on many contingent congruent features that are only formed through this particular manifestation of the tendency for congruence.

As a general overview of the model and rules, the butterfly schema uses functional chords in a single key, and a particular metrical structure (congruent at the level of regular functional harmonic change and at two immediately higher levels, shown in the dot structure), which correspond with the three main surface features. The T–Dep…Ret–T satisfies the permissible distance range in the Dep and Ret regions (δ = 5–8), constituting variably congruent chords, i.e., V, ii/ii^{o}, iii/III, or VI/vi. In the Dep and Ret regions, functional chords close to the tonic

($\delta = \sim0$) are not permissible, and chord distances greater than $\delta = 8$ must resolve to legal functions. In the T regions, functional tonic chords ($\delta = 0$) must be used (and less stable chords in these regions must resolve to the T functions). For the schema to be multiparametrically congruent, it must include textural grouping at the level of regular functional harmonic change and at two immediately higher levels, together with a maximally congruent 1:1 harmonic-rhythm ratio, or a minimally congruent 3:1 ratio.

As discussed, when minimally and maximally uniparametrically congruent features occur together, the schema is classified as minimally congruent. A schema is only maximally multiparametrically congruent if *all* features are maximally uniparametrically congruent. The third region of the butterfly has been shown to be important for defining maximal UC and MC because it is in a position where the textural grouping, harmonic-rhythm ratio, and metrical structure have the potential for significant interaction. A chord V in the third region provides greatest UC in the chord progression and possibly MC overall (i.e., I–Dep…V–I) because this is where textural grouping and metrical structure are also potentially maximally UC. For an overall maximally multiparametrically congruent butterfly schema, textural grouping must also occur at the level of regular functional harmonic change and at two immediately higher levels, and the harmonic-rhythm ratio must be 1:1. The opening of Beethoven's Piano Sonata Op. 2 No. 3, i, in Figure 5.21, is a maximally multiparametrically congruent butterfly schema, because all features are maximally uniparametrically congruent. While there is the potential in grammars for an admixture of congruent and noncongruent features, the butterfly schema model is formally classified into minimally and maximally multiparametrically congruent schemata because this is necessary for a quantitative analysis:

(*xiv*) *A minimally congruent butterfly schema =*
(*rule x*) + ((*rule v* + *rule xiii*) ***or*** (*rule v* + *rule xii)* ***or*** (*rule vi* + *rule xiii*)).

(*xv*) *A maximally congruent butterfly schema = rule x + rule vi + rule xii.*

5.2 A Survey of Butterfly Schemata (*c.* 1750–*c.* 1850)

As elaborated in Sections 1.3.1 and 4.1.2, a non-inductive and rational account of musical cognition is assumed in this thesis (although not explored as a focus). Congruence and noncongruence varies in musical structure because while it is universally constrained (by cognition), it is also particular, being emergent and constrained by culture. The argument put forward was that structure forms with a spectrum of emergent interacting features, and since these are reconstructed in cognition, no *a priori* mental tools or heuristics could inductively infer these novel arrangements. Rather, a rational, abductive, and intuitive mind must interpret the various degrees of congruence and noncongruence therein. Indeed, an 'inductive' computational algorithm could be designed to carry out automated corpus analyses (similar to those reviewed in Section 3.1) of multiparametrically congruent schemata. However, this would not serve the purpose of showing that music is inductively inferred since a schematic category must be predefined by an analyst (or computational musicologist) prior to a corpus analysis. Thus, the butterfly schema, which represents a multiparametrically congruent category, is not fully determined by cognitive principles, nor can its inference from musical structure be necessary — although principles of cognition indeed influence musical structure.

It follows from this that for the abstraction of relevant rules, empirical surveys must be conducted with a knowledge of the particular context. It has been well-documented that in statistical analyses and corpus studies of music, a knowledge of context is required to understand specific data (e.g., Gjerdingen, 2014). Highlighting the importance of context, Ito (2014) presents a corpus analysis of Mozart examining the validity of Heinrich Koch's theory on the metrical placement of cadences, that cadences generally take place on strong beats of metre. Ito (2014) finds that Koch's theory applies strongly to pieces in common time (cadences generally take place on strong beats) but less so to pieces in various other time signatures. Also, Prince and Schmuckler (2014) conduct a corpus analysis examining the congruent interaction between tonal and metrical structure, arguing that while there is general stability between these systems over composers, time signatures, and modes, there is some variation in the relationship with certain composers. That is, 'metrically stable positions in the measure preferentially feature tonally stable pitches' (Prince and Schmuckler, p. 254), but how this preference is realised varies. With some modern composers this principle of congruence is less pronounced, and with other composers, of different periods, the correspondence between these structures is more subtle:

> […] Bach placed tonally stable pitches in multiple places throughout the measure, while reserving metrically stable positions for tonally stable pitches. Conversely, Chopin was more likely to limit tonally stable pitches to metrically stable points, but was less restrictive on what pitch classes occurred at these times (i.e., tonally unstable pitches could also occur at strong metrical locations). (Prince and Schmuckler, p. 265)

These corpus studies are noteworthy not only because they support the present notion of a tendency for congruence between features of different parameters, but because they show that context (i.e., the influence of cognition and culture) is an important inclusion which must be understood through the critical engagement and abductive intuition of the analyst.

A survey was carried out to determine whether butterfly schemata are more common in the Classical period (*c.*1750–*c.* 1800) than the Romantic period (*c.*1800–*c.* 1850). This was necessary to support the claim that the particular manifestation of the tendency for congruence (quantified through the rule of MC) in the Classical grammar causes butterfly schemata (Section 1.4). The sample comprised 973 movements and single-movement works from ten genres of five European composers of instrumental music, during the period *c.* 1750–*c.* 1850. (The movements and pieces examined in the survey are listed in Appendix A.)

The survey develops the methodologies of associative-statistical theories (e.g., Gjerdingen (1988) and Byros (2009a)). A Romantic-period sample was chosen over a Baroque-period sample or Renaissance-period sample for comparison with the Classical period because a comparison with the earlier grammars was thought to be, to some extent, less illuminative. Butterfly schemata could not feasibly form with any consistency in Baroque grammars owing to their predominant textural grouping at lower levels of the metrical hierarchy (generally the 2nd and 3rd levels). Romantic grammars are more akin to those of the Classical period in textural and metrical structure, but not in chord progression. Instrumental art music of the Romantic period often comprises chords with large distances from the tonic in diatonic space, or employs chromatic basic spaces, such as octatonic or hexatonic spaces (Lerdahl, 2001), which limit the frequency of the specific type of MC that constitutes butterfly schemata. While the tonal language of Romantic-period grammars generally does not have the necessary type of MC to consistently produce butterfly schemata, there might nonetheless be a low output of butterfly schemata generated therein, and so a difference in output between those and the Classical grammar would verify subtle but significant variances between the two grammars.

It was necessary to incorporate predefined corpora for an unbiased comparison of works in the respective grammars. A complete instrumental genre of a single composer provided an ideal way of demarcating a sample because these genres were originally grouped this way by the composer, eliminating any preferential bias of the analyst. Using works of prominent composers was also an objective way to elect corpora, avoiding cherry-picking of data to suit hypotheses. (Appendix B presents charts showing the distribution of schemata for each composer and genre.)

The corpora span both periods, comprising two instrumental genres from each of five composers: Haydn string quartets and piano sonatas, Mozart symphonies and string quartets, Beethoven string quartets and piano sonatas, Schubert string quartets and piano sonatas, and Chopin mazurkas and nocturnes. The survey recorded the quantity of butterfly schemata present, the type of congruent features found, and the voice-leading schemata that were embedded in the butterfly schemata. The maximum quantity of schemata that were recorded in a movement or single-movement piece was limited to one in order to avoid recording the exact and inexact repetition of schematic material. Table 5.1 shows the quantity and type of butterfly schemata found.

c. 1750–*c.* 1800					*c.* 1800–*c.* 1850				
	Total Movts.	**Movts. With Butterfly**	**Max. MC**	**Butterfly With Meyer**		**Total Movts.**	**Movts. With Butterfly**	**Max. MC**	**Butterfly With Meyer**
Haydn Quartets	220	53 (24%)	48	7	**(Post-1800) Beethoven Sonatas**	60	7 (12%)	6	2
Haydn Sonatas	144	22 (15%)	18	5	**(Post-1800) Beethoven Quartets**	59	4 (7%)	4	1
Mozart Symphonies	146	38 (26%)	25	8	**Schubert Quartets**	53	2 (4%)	1	0
Mozart Quartets	86	17 (20%)	10	4	**Schubert Sonatas**	67	8 (12%)	6	0
(Pre-1800) Beethoven Sonatas	48	10 (21%)	8	1	**Chopin Mazurkas**	56	3 (5%)	2	0
(Pre-1800) Beethoven Quartets	12	3 (25%)	1	0	**Chopin Nocturnes**	22	0 (0%)	0	0
Total	**656**	**143 (22%)**	**110**	**25**	**Total**	**317**	**24 (7.6%)**	**19**	**3**

Table 5.1 The quantity and types of butterfly schemata found in samples of movements and single-movement works, *c.* 1750–*c.* 1850.

Comparing the data in Table 5.1, 22% of movements in the Classical-period samples contain butterfly schemata, compared with 7.6% of movements and single-movement works in the Romantic-period samples. This means that butterfly schemata are almost three times more prevalent in the Classical-period samples than in the Romantic-period samples. Since butterfly schemata are multiparametrically congruent structures, these findings suggest that the rule of MC is a constraint on the formation of these schemata, and is more commonly manifest in this particular form in the Classical grammar than in the Romantic grammar.

An independent-samples t-test was performed to see if the difference between the means of the Classical-period samples and Romantic-period samples are statistically significant (presented in Appendix C). The test compares the variability between samples against the variability within samples. The test showed that the difference between the means of the samples of each period was very significant (*t* = 6.10; *df* = 10; *p* = 0.0001 (two-tailed distribution)), establishing that these data are unlikely to have occurred through chance.

The degree of congruence within individual features of butterfly schemata also indicates the influence of the rule of MC. As pointed out, in the present formulation, a minimally multiparametrically congruent butterfly schema can contain an admixture of minimally and maximally uniparametrically congruent features, but a maximally multiparametrically congruent butterfly schema must have all features maximally uniparametrically congruent. Of the butterfly schemata found in the Classical period samples, 76% were maximally multiparametrically congruent schemata.[37] The remaining 24% of butterfly schemata therefore had minimally uniparametrically congruent features or an admixture of minimally and maximally uniparamterically congruent features (categorised as minimally multiparametrically congruent schemata). That is, in the Classical period samples, butterfly schemata with entirely maximal uniparametrically congruent features (1:1 / T–Dep…V–T) are about three times more common than butterfly schemata with minimally uniparametrically congruent features (3:1 or 2:1 / T–Dep…Ret–T) or an admixture of minimally and maximally uniparametrically congruent features. These findings support the claim that the tendency for congruence generates butterfly schemata, since congruence is consistently manifest in and between the particular features of the Classical grammar.

The survey could not compare equal quantities of works in the Classical- and Romantic-period grammars due to the general trend for greater quantities of works in genres of the former grammar. This means it was necessary to compare percentages of schemata found in each period, rather than absolute values. It could be that a larger sample, with a wider selection of instrumental composers and genres (perhaps representing less salient composers), may produce different results to the present findings. However, the wide variance between the quantity and type of butterfly schemata found in the Classical and Romantic samples

[37] Compiling the results in Table 5.1 it can be seen that 80% of butterfly schemata were maximally multiparametrically congruent in the Romantic period. This means that while butterfly schemata are much less popular during the Romantic period, the type of grammatical structuring that produces butterfly schemata, even though generally less common in this period, is present in a residual form in the Romantic period.

surveyed suggests that alternative samples would not have produced significantly different results.

Gjerdingen's Meyer, here formulated as a multiparametrically congruent child schema of the butterfly schema, was the most common voice-leading variant, but occupies a small fraction of the butterfly schemata found in the Classical samples. Indeed, the Meyer child schema was present in only 17% of the butterfly schemata in the Classical samples. As noted, butterfly schemata were present in 22% of movements in the Classical samples, which means that the Meyer was found in only 3.7% of Classical movements overall. Its sparse population gives reason to question the Meyer (as formulated in terms of the strict multiparametrically congruent criteria described here) as a unit of style during the Classical period. However, if Gjerdingen's associative methodology was employed, perhaps a greater quantity of Meyer schemata would have been identified in the samples — yet, the schema theory paradigm does not provide a framework which systematically relates schema features (see Section 4.3). The present grammatically congruent approach is preferable in this respect because choices are empirically grounded, without relying on interpretative judgments. The present approach does not absolutely diminish any possible significance of voice-leading variants as part of the late-Classical instrumental grammar, as they still have meaning when viewed from particular historically situated listening perspectives. Moreover, voice-leading features might be more consequential in the early Classical period (*c*. 1700–*c*. 1750), as suggested by the flourishing *solfeggi* and *partimenti* practice during this period (see Sanguinetti, 2012).

5.3 Top-Down Hierarchical Selection

This chapter has thus far accrued evidence that the tendency for congruence, in a particular manifest form in the Classical grammar — quantified through the rule of MC — generates butterfly schemata. This section shows that, in addition to the tendency for congruence, the Classical grammar and butterfly schema are a product of the top-down hierarchical selection (HS) of their features, which is driven by culture.

Top-down HS draws from selectionist frameworks, such as those of evolutionary biology, memetics, and behavioural psychology. In top-down HS, features are cumulatively selected against each other in a particular order, forming a complex hierarchical structure. The order in which features are culturally selected against each other, their order of dependency, is revealed in the top-down HS model. Furthermore, the interaction between the universal tendency for congruence (which is a product of cognition) and the selective mechanism of culture is revealed.

5.3.1 The Tendency for Congruence and Hierarchical Selection

Universal constraints of psychology — reducible to the tendency for congruence — interact with the particular constraints of grammars. These particular, yet universally constrained features, are culturally selected in grammars, which are situated in distinct histories and geographies. The universal and the particular do not exist as *entities* in grammars and schemata, but rather they are *constraints* on musical structure (see Section 4.1.1).[38] Therefore both these notions are challenging to define and tease apart because both interact within the same material phenomena, the musical structure itself. However, it is possible to separate these through observations and comparisons of musical grammars and the application of the tendency for congruence.

Jan (2007, p. 141) explores Meyer's observation on the use of 6/4 progressions, which frames the problem of distinguishing between universal and particular phenomena:

> [A]lthough *nature*—that is, the constraints of human cognition—establishes the conditions that make it possible for the cadential $^6_4$ progression [I^6_4–V–I] to imply and define a tonal center, the frequency with which the progression was replicated (chosen by composers) depended in large measure on cultural constraints—that is, on *nurture*. (Meyer (1992, p. 490), quoted in Jan (2007, p. 141))
>
> In considering memetic evolution it is necessary to understand which elements of a cultural-evolutionary system are invariant and which are mutable; that is, it is important to be clear about the distinction between nature and nurture in memetic evolution. One might say that processes operating on Meyer's level of laws determine the environment for memetic propagation, and then the

38 The term 'constraint' is generally analogous to preference rule. 'Constraint' is used in *TPS* to show the most cognitively preferred pitch-space relationships.

> outcomes of the selection processes which operate within the terms of these laws determine the complexion of the rules (including systems of organization) which arise. (Jan, 2007, p. 141)

According to the tendency for congruence, cognition engenders a propensity for regular, duple, and hierarchical rhythmic, harmonic, and metrical structure, using chords with close distances in pitch space from the tonic. While divergences from the tendency for congruence are conceivable, because culture is infinite (Lerdahl, 2001, p. 381), the Classical grammar tends to generate particular congruent structuring, found in butterfly schemata. However, while adhering to the rule of MC, the butterfly schema contains features that are cumulatively selected, which means that universal constraints interact with particular constraints. As discussed in Sections 1.2.2 and 4.1.1, this means that even though the universal constraints generally limit the type of structuring that tends to form, culture still has to select these preferences out of possible alternatives. This nuances the relationship between the universal and particular, since the latter is often formed from the former.

The tendency for congruence follows Humboldt's (1836, 93) notion (referring to language grammar) that cognition regenerates (*wiedererzeugt*) similar types of structures. In the domain of pitch, the diatonic basic space is a uniparametrically congruent construct that is presumably regenerated by cognition, and follows the psychoacoustic 'constraints on basic spaces' (Lerdahl, 2001, pp. 272–274). The diatonic basic space thereafter constrains the chord progressions of the Classical grammar. Thus the chord progression of the butterfly schema closely follows the universal cognitive constraints on basic spaces (in the pitch-space theory of Lerdahl (2001)). While this chordal schema is particular to the Classical grammar, many grammars can feasibly have a similar schema, because they form from universal constraints, and so are not necessarily unique.

Lerdahl (2001, pp. 268–274) theorises that pitch systems of all cultures adhere to the same universal constraints, described as the 'constraints on basic spaces' in *TPS*. The constraints on basic spaces are sensory consonance, almost-halving, uniqueness, maximal evenness, two step sizes, two generators, and small steps (Lerdahl, 2001, pp. 272–274) (examined in Section 2.3.7), which are elements of uniparametrically congruent pitch systems. The most commonly used basic spaces in grammars and schemata follow these constraints, such as the diatonic, pentatonic, octatonic, hexatonic, and whole-tone spaces (Lerdahl, 2001, p. 273). In support of this theoretical claim, Krumhansl (1990) contends that listeners of all cultures represent pitch distances in distinct systems in a broadly similar way, providing they have sufficient exposure to the particular musical system. While these universal constraints act on all grammars, they are incorporated by degree, and in various arbitrary forms. The Classical grammar generally uses diatonic space, which is the most preferred basic space in terms of the constraints on basic spaces (Lerdahl, 2001, pp. 268–274). Lerdahl (2001, p. 190) argues that '[t]he [diatonic] basic space is … a specific instantiation of general principles of organization'. Thus, basic spaces tend to form according to the constraints on basic spaces, which accords with the present notion of the tendency for congruence.

Cognition also regenerates universal phenomena in the domain of rhythm, and these phenomena are likewise subject to cultural selection. The tendency for congruence provides a constraint for duple regular hierarchy of metrical structure, which governs grammars. However, grammars are free to manifest these tendencies in arbitrary arrangements, due to the action of culture. Indeed, metrical structure is a universal and particular phenomenon because it is an *a priori* concept that originates in cognition (London (2004), but is also formed through the particular constraints of culture. Both these factors cause the metrical structures of grammars. Separating universal (cognitive) constraints from particular (cultural) constraints — in the domains of rhythmic structure and metrical structure — is problematic because these phenomena are, similar to the domain of pitch, locked together in musical structure. Indeed, since the inference of metrical structure in real-time is underdetermined by the surface structure, being abstracted from mere surface-level cues, metrical structure must be at least partly an *a priori* concept. However, metrical structure is also culturally conditioned, as evinced by the variety of particular metrical types found in grammars (see Section 5.1.2).

In the metrical domain, while music often exploits the cognitive preference for regular, duple hierarchical constructions, metrical structure is also a product of actual accentuation patterns, which can be diverse (both in terms of cultural practice and in terms of the temporal unfolding of pieces). Therefore, although metrical structure is limited by cognitive preferences (governed by the tendency for congruence), it admits variability and plurality of structure due to the influence of culture. In the Classical grammar, metrical structure forms a regular duple hierarchy at the level of regular functional harmonic change and at two immediately higher levels, which is a culturally conditioned manifestation of the universal constraint for the tendency for congruence.

Well-formedness of tonal and metrical structure, as espoused in generative theories, are essentialist conceptions. The systemic diversity and variability in tonal and metrical structures of the Baroque, Classical, and Romantic grammars, outlined in Section 5.1.2, illustrate this point. Well-formedness of rhythm, metre, and pitch, are not incontrovertible structures of Western music. The particular manifestation of the tendency for congruence in metrical and tonal structures is memetically transmitted in the Classical grammar. These structures thereafter condition less abstract features that form in this grammar through a top-down selection hierarchy. Thus, top-down selection hierarchies show how the tendency for congruence and culturally conditioned features are interdependent in grammars and schemata.

5.3.2 Top-Down Hierarchical Selection

Figure 5.25 is an abstract representation of a causal, or selective hierarchical system. In hierarchical systems, the normative order of causation or selection is bottom-up, from part to whole. That is, smaller parts are grouped into larger wholes. In bottom-up selection, smaller parts are the cause of larger parts (Ellis, 2005, 2008), or select, larger parts (Okasha, 2012).

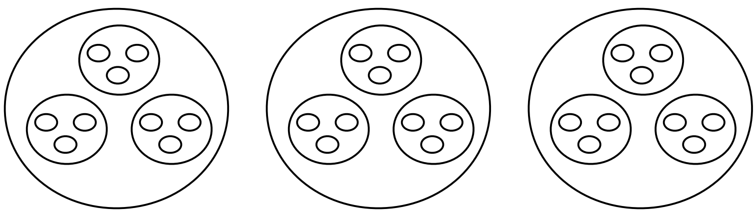


Figure 5.25 An abstract representation of a hierarchical system (Okasha, 2012, 41).

Bottom-up selection describes the process of real-time listening (explained in Section 4.1.2), where smaller, more concrete parts of the musical surface are abstracted into larger wholes (inside the head). That is, in bottom-up selection, higher-level features are caused by lower-level features. However, more broadly in musical structure, outside the head, features at higher levels either supervene on lower-level features, meaning that they are governed by lower-level features, or, conversely, emerge from lower-level features, which means they are irreducible to them. (Emergence can be used to describe the introduction of novel structure or paradigm that originates from, but is irreducible to, the rules or theories of another structure or paradigm (Ellis, 2008, 2005). For example, biology is irreducible to chemistry, and chemistry is irreducible to physics.) Real-time listening is necessarily bottom-up, owing to the particulate nature of music perception and cognition. In real-time listening (which occurs inside the head), higher-level structuring generally supervenes on lower-level structuring, such as where a melody supervenes on its note events and where harmony supervenes on the melodies that comprise it.

Top-down selection and emergence, where large wholes govern the organisation of parts, operates in various complex systems of the world (Ellis, 2005, 2008; Okasha, 2012, p. 49). This occurs in musical structure through the action of culture, which acts from top-down, from whole to part. That is, in the top-down selection hierarchies of culture, higher-levels of structure are often emergent; they are not necessarily governed by lower-level features. Top-down selection is incorporated into systems theory (i.e., von Bertalanffy (1968)), the philosophy of biology (Okasha, 2006, 2012), and causal systems (Ellis, 2005, 2008). In biology, selection can take place at the group, species, individual, or gene levels (Okasha, 2006). Top-down selection in biology would be that the species level would govern survival at the group level, which would, in turn, govern the survival of individuals. However, music may be similar to many other complex systems, where bottom-up and top-down causation (or selection) work simultaneously. Musical structure probably involves multiple causality or multi-level selection, which might explain many complex systems. Ellis (2005) describes multiple causal processes in real-world phenomena. Ellis (2005) uses bottom-up, same-level, and top-down causality to explain aircrafts in flight. The bottom-up explanation is physical (aircrafts fly due to the air pressure difference under the wing, causing lift). The middle-level explanation is logistics (aircrafts fly because air traffic controllers and pilots coordinate their flight paths). The top-down explanation is design (aircrafts fly because they have been engineered to do so). All these levels must interact simultaneously for an aircraft to fly.

Top-down HS in the butterfly schema and Classical grammar is generally governed by history and culture, where higher levels of structure select lower levels of structure, from whole to part. In the present use of top-down selection, although higher-level features constrain the arrangement of lower-level features, they do not change the internal organisation of lower-level features. This differs from some explanations of top-down causation and selection, where higher levels are thought to change the internal structure of lower-level phenomena — which is a more complex conception (explained in Okasha (2012, p. 49)). In the butterfly schema and Classical grammar, the top-down selective mechanism of culture merely organises lower-level features. The fittest features (on account of their MC and correspondence with features at higher levels) are selected at each level of the hierarchy, from top down, producing a culturally constrained architecture. Less fit features (which are noncongruent and do not match the higher-level features) can potentially form at any level of grammar, but are less common at higher levels, tending to form at lower levels where selection pressures are weaker. The top-down HS of features by culture is a more fundamental constraint on musical structure than those imposed by the actions of individual composers or intraopus styles. Novel or creative phenomena, such as those introduced by composers, are selected out of the selection hierarchy, at each level, because they are not multiparametrically congruent or do not match higher-level features. (To reiterate, top-down HS in culture occurs in the converse causal direction to the bottom-up cause of musical structure, which is the perception and cognition of musical structure in real-time listening.)

5.3.3 The Mechanism of Selection in Memetics and Behaviourism

The evolutionary 'algorithm' of variation, selection, and replication (Jan, 2007, p. 8) governs the formation of grammars and schemata. Variation is implicit in the top-down HS model, but selection and replication (heredity) are more explicitly described. The present theory of top-down HS is based on hierarchy theories in Meyer (1989), Jan (2001, 2007), Ellis (2005, 2008) Okasha (2006, 2012), as well as the notion of hierarchical systems in systems theory (Bertalanffy, 1968). Jan (2001, 2007, 2013) proposes that the same evolutionary processes that operate in genes analogise to music (following the concept of 'universal Darwinism', Dawkins, 1983b). Memes (which are equivalent to features, since they are distinctive attributes in grammars and schemata) are selected against each other, and also form memeplexal relationships with other memes. Memeplexes (equivalent to schemata) comprise memes that combine for mutual benefit, competing against their alleles for particular loci in structure.

The analogies of memeplexes and alleles to features are useful for the present view of selection because, in broad terms, features are selected, work together, or compete with each other at particular levels in selection hierarchies. Memes are proposed to occupy particular niches, which are selected against each other hierarchically. Selection is also an important process in memetics. '[M]emes, selected against the background of each other, "cooperate" in mutually supportive memeplexes — supportive within the memeplex but hostile to rival memeplexes' (Dawkins, in Blackmore, 1999, p. xv). Jan's (2013) notion of musical memeplexes in galant schemata also supports the present concept of top-down HS. Features combine with each other in parallel memeplexes, according to their relative fitness. In support of an analogy between genes and memes, Dawkins points out how genes are selected against each other in the gene pool:

> Each gene is seen as pursuing its own self-interested agenda against the background of the other genes in the gene pool—the set of candidates for sexual shuffling within a species. Those other genes are part of the environment in which each gene survives, in the same way as the weather, predators and prey, supporting vegetation and soil bacteria are parts of the environment. From each gene's point of view, the 'background' genes are those with which it shares bodies in its journey down the generations. (Dawkins, 1976/1989, p. ix)

The top-down HS model of the butterfly schema can thus be viewed as a giant memeplex, a complex of features that has been cumulatively selected through time. Thus 'memes build on memes' (Blackmore, 1999, p. 78) in parallel, resulting in a constrained architecture of features. The particular order of selection of these memes in the butterfly schema and Classical grammar, the order of dependency, is revealed in the top-down HS model. Jan

(2007, pp. 176–177), after Dawkins (1986, pp. 45–51), shows that cumulative models are the most plausible explanation of cultural complexity. Memes build on existing memes to produce complexity. Jan compares single-step evolution with the accumulation of cultural features, arguing that,

> [i]t is clear what would happen if such a collection of memes [referring to structural memes] were subject to the single-step mechanism Dawkins first employed in his experiment. On this model, every composer would start afresh, with a *tabula rasa*, assembling memes to make phrases without reference to the attempts of earlier composers. One can see how difficult it would be to move to such a 'target' configuration as that represented by sonata form. To estimate the odds involved, we would need to determine the number of events necessary to give rise to such a form. If we decide, conservatively, that twenty specific characteristics are necessary to appear in a particular order for a given movement to belong to the class of sonata-form movements (to confom to what Carter terms the 'sonata principle' (1987: 89)), then, using Dawkin's equation, the statistical probability of a Josquin or a Palestrina spontaneously generating a sonata-form movement are $(1/20)^{20}$, where for each of the twenty 'slots' there are classes each comprising thousands of alleles of analogous configuration, function or structure suitable to occupy it. (Jan, 2007, pp. 178–179)

The top-down HS model also draws from the mechanism of selection espoused in behaviourism (Skinner, 1981), discussed in Section 3.2.3. The 'selection by consequences' of Skinner unites natural selection, operant conditioning, and cultural selection. Skinner (1981, p. 503) argues that '[o]nly past consequences figure in selection' in these systems. The notion of selection by consequences can be applied to the top-down HS of features of the butterfly schema. The existence of a feature in a particular time and place is not a consequence of it being an aesthetically appropriate, artistic, or an intellectually significant addition to a grammar or schema. Lower-level features are merely similar features to those elicited by higher-level features that have shaped and maintained those features in the past. For example, the lower-level feature of a major chord is not necessarily an intentionally included object of musical structure, but a consequence of a selection history, in particular, the higher-level feature of tonality, which has shaped and maintained it in the past. Selection by consequences is confusing because it is the reverse of what is generally thought to be true (Baum, 1990). Features of grammars and schemata are generally considered intentionally devised attributes of musical structure that are products of rational composers. However, this is controverted in the top-down HS model because culture is shown to be a stronger influence on grammars than intelligent design. The mechanism of top-down HS can now be outlined specifically for the butterfly schema and the Classical grammar, while incorporating the psychological tendency for congruence (formalised through the rule of MC):

(*xvi*) *Features of the butterfly schema and the Classical grammar are culturally selected against each other in a top-down hierarchical order, with the rule of MC governing each level of the hierarchy*.

As discussed, top-down HS suggests that composers did not freely choose features in music of the Classical grammar and butterfly schemata, but features survive at each level of the hierarchy, and then act as further filters for continued selection. Musical structure is generally not the product of compositionally aware creators, but is culturally selected. Thus, in grammars, the order of the dependency of features is pre-selected, and so composers are constrained by this orthodoxy. In the past, these features satisfied the universal (rule of MC) and cultural constraints that have filtered them. Since top-down HS governs grammars of a particular time and period which composers shared and rarely deviated from, it is unlikely and perhaps inconsequential whether they are partly rationally constructed. Individual and intraopus styles are elusive and cannot be easily distinguished from particular features, or the universal rule of MC. Whether it is a composer's choice to employ grammatical structure is perhaps not a meaningful question if they do not deviate from the top-down selection hierarchy of the grammar. The creative elements of pieces are, by definition, not generally repeated — they are emergent — and so do not form part of the grammatical or schematic structure.

While the top-down HS model is proposed to characterise instrumental music of the Classical grammar, it should be noted that the arrangement of features is not fixed in time, since grammars can evolve through historical periods. Features change over time, and so too does the order of dependency between features. Higher-level features are more stable than lower-level features since they manifest more abstract structure that is resistant to change. Features at lower levels therefore can become extinct, or used for different functions. A change in the application of a feature, rather than the presence or absence of features, is similar to the concept of a change of strategy, described by Meyer (1989, p. 20) (discussed in Section 3.2.5). An example of a change in strategy can be seen in the transformation from contrapuntal to homophonic grammars that occurred in the seventeenth century. During this time, the parsimony of voice leading, which broadly dictated harmonic structure in contrapuntal grammars (such as those of the Renaissance and Baroque periods), gradually became dominated by functional harmony (Berry, 1976; Swain, 2002, p. 68).

5.3.4 Determining the Order of Selection in the Classical Grammar and Butterfly Schema

As discussed, the top-down HS model of the Classical grammar and butterfly schema shows that higher-level features select lower-level features. The rule of MC governs the rhythmic and metrical features, and acts on all levels of the hierarchy. The rule of MC selects for duple regular hierarchy and close pitch-space distances from the tonic. The particular grammatical features of the Classical grammar, which are chord progressions close to the tonic, textural grouping (at the level of regular functional harmonic change and at two immediately higher levels), and simple harmonic-rhythm ratios, are selected by these higher-level constraints. Multiparametrically congruent local schemata, such as the butterfly schema, select highly concrete and particular voice-leading schemata. The butterfly schema admits a large degree of flexibility in the possible voice-leading structures (such as those of Gjerdingen (1988, 2007)) that can occupy its chord progressions. Indeed, the butterfly is a more general pattern than the voice-leading schemata of Gjerdingen (1988, 2007), as shown in the survey in Section 5.2.

The order of dependency of features in the Classical grammar and butterfly schema challenges the orthodoxy on the relationships between features in many generative theories. In *GTTM*, the order of dependency of features broadly runs from grouping structure, metrical structure, time-span structure, to prolongational structure (although there are feedback loops that can change this order). It is questionable whether the order of dependency in *GTTM* reflects all tonal grammars (or all grammars of the world). A variety of different orderings are possible because any feature of one parameter can conceivably inform any feature of a different parameter. Particular grammars determine the order of dependency between features in its respective schemata.

In *GTTM*, accentuation in grouping structure is proposed to give rise to metrical structure. However, the top-down HS model of the butterfly schema suggests that *Gestalten* in melodic phrase grouping are not the cause of metrical structures. Actually, the opposite is the case in the Classical grammar and butterfly schema, metrical structure causes grouping structure in history and culture (outside the head). Moreover, melodic grouping is not as significant as textural grouping for creating metrical structure. Textural grouping cues metrical structure in real-time listening but is selected by metrical structure in culture. Therefore in the top-down HS model, textural grouping is shown to be culturally selected by the particular uniparametrically congruent metrical structure. This particular manifestation of the universal constraint for regularity and hierarchy in metrical structure (a product of the tendency for congruence) is selected by culture. Indeed, the chord progression, textural grouping, and harmonic-rhythm ratio of the Classical grammar are selected by a particular memetically transmitted metrical structure. This top-down determination of musical structure might resemble the order of dependency of other grammars of Western music, but further investigation is required to demonstrate a wider significance.

5.3.5 A Top-Down Hierarchical Selection Model of the Butterfly Schema

The discussion here presents a top-down HS model of the butterfly schema, showing the causal mechanism of culture that explains how the features of the Classical grammar and butterfly schemata form. The butterfly schema is a structure that is locally selected by the multiparametrically congruent features of the Classical grammar. Figure 5.26 shows the top-down HS model of the butterfly schema. Higher-level features cumulatively select lower-level features. All features are ultimately selected by the rule of MC, which acts on all features of the hierarchy.

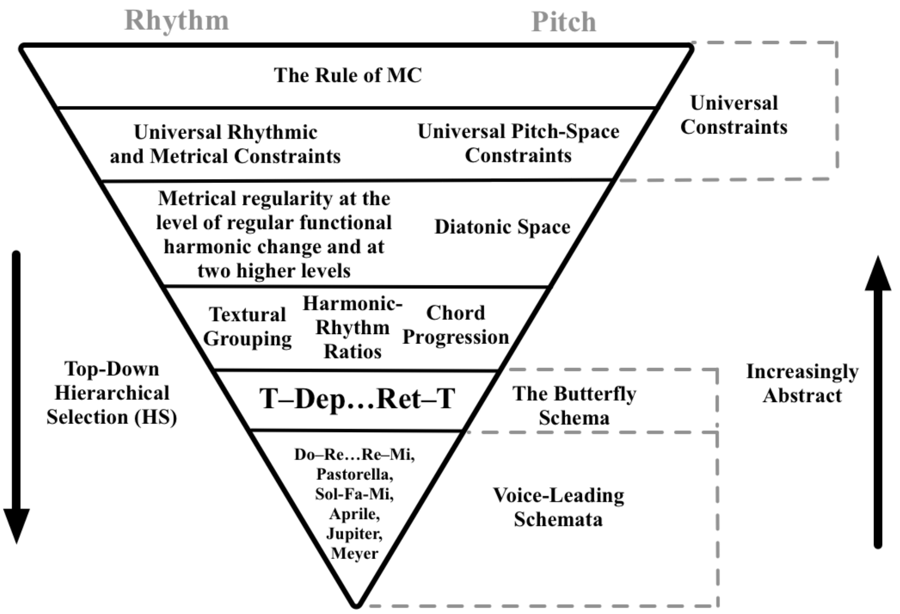


Figure 5.26 The top-down hierarchical selection model of the butterfly schema.

The hierarchy comprises two interdependent domains, rhythm and pitch. The top level of the hierarchy contains the rule of MC, which constrains all levels of the hierarchy. The rule of MC configures the universal constraints in the domains of rhythm and metrical structure (regular, duple, and hierarchical structuring) and pitch (closeness to the tonic in pitch space). These higher levels are the most abstract, whereas features become increasingly concrete towards the bottom of the hierarchy. The level below the universal rhythmic, metrical, and pitch-space constraints is the level that is, in an important sense, particular to the Classical grammar (chosen out of possible alternatives), where diatonic space and metrical regularity (at the level of regular functional harmonic change and at two immediately higher levels) are selected. At the level below this, textural grouping, harmonic-rhythm ratio, and chord progression are selected, which are likewise particular to the Classical grammar. The butterfly schema is selected at the next level down, at the local schema level. The butterfly schema is the parent schema of a number of Gjerdingen's (2007) voice-leading child schemata (the Do–Re…Re–Mi, Pastorella, Sol-Fa-Mi, Aprile, Jupiter, and Meyer).

As discussed, divergent (noncongruent and incompatible) features can potentially form at each level in Classical instrumental music, but tend to be selected out, since the top-down HS and the rule of MC shown in Figure 5.26 is normative in this grammar. Various other schemata in the grammar can form at the level of local schemata in this model (although not shown in Figure 5.26) which would be multiparametrically congruent providing they are supported by the same higher-level features described, representing different schemata. Noncongruent and incompatible features might occasionally form at the local schema level and below in Classical instrumental music, rather than at higher levels, which are more stable. However, the formation of noncongruent and incompatible schemata, even at lower levels, such as the local schema level, is rare and so does not form part of a grammatical hierarchy. An examination of consistent noncongruence in top-down HS in grammars would be interesting because it would indicate the counteraction of the selective mechanism of culture against the universal tendency for congruence. However, such noncongruent schemata are relatively sparse and are not well documented, perhaps with the exception of Byros' (2009a) *le–sol–fi–sol* schema. By contrast, the butterfly schema is common in the Classical grammar and formed from highly congruent features. Culture can potentially override the universal constraint of the tendency for congruence, selecting arbitrary features, but arbitrary structuring is rare and so does not diminish the generality of this constraint.

Top-down HS is a useful mechanism for illustrating the interaction between features of grammars and schemata which are selected by culture but also constrained by the tendency for congruence. Moreover, it provides a system that permits culture to conflict with the tendency for congruence, although noncongruent structuring tends not to occur, especially in the Classical grammar. Top-down HS is an explanation of the cultural cause of features in the Classical grammar, based on the variation, replication, and selection of features. More complex relationships between features in other grammars might place limitations on the generality of top-down HS. Nonetheless, it provides a stimulating hypothesis about the interaction between the cognitively produced tendency for congruence and the cultural constraints in grammars and schemata.

5.4 Conclusions

This chapter quantifies the uniparametrically congruent features of the butterfly schema — the chord progression, textural grouping, and harmonic rhythm — through separate formulations that are thereafter combined into a multiparametrically congruent model. The uniparametrically congruent chord progression is expounded with reference to Lerdahl's (2001) pitch-space theory. Chord functions, which modify Lerdahl's generalised functions, depict close distances from the tonic in pitch space. Two forms of the butterfly schema progression are presented, namely, minimally congruent (I–Dep…Ret–T), and maximally congruent (I–Dep…V–I) versions. Minimally congruent butterfly schema progressions must have Dep and Ret functions in the pitch-space distance range, δ = 5–8. Maximally congruent progressions are more constrained. In this form, while the Dep function is in the range, δ = 5–8, the Ret function must be a chord V (δ = 5).

Textural grouping at the level of regular functional harmonic change and at two immediately higher levels of metrical structure is a uniparametrically congruent feature of the Classical grammar and butterfly schema. Textural grouping builds on the notions of textural accent (Lester, 1986), fusion (Lerdahl and Jackendoff 1983), and contrapuntal grouping (Lerdahl, 2001; Temperley, 2001), but occurs between distinct levels of texture. Textural grouping is a more significant feature of certain grammars and schemata than has been previously recognised, and is here argued to be a primary influence on metrical structure. Textural grouping varies depending on the grammar, contrasting with the generalised portrayal of texture and well-formed metre in generative theories (e.g., Lerdahl and Jackendoff (1983), Lerdahl (2001), Temperley (2001), London (2004)). As shown above, textural grouping occurs at low levels of metrical structure in Baroque grammars.

Simple and regular harmonic-rhythm ratios are uniparametrically congruent features of the butterfly schema and the Classical grammar, permitting regular and deep duple hierarchies that enable MC. Harmonic-rhythm ratios are of variable UC. The 1:1 harmonic-rhythm ratio is more uniform and permits greater duple hierarchical structuring than the 3:1 ratio. The combined uniparametrically congruent features of chord progression, textural grouping, and harmonic-rhythm ratio result in minimally and maximally multiparametrically congruent butterfly schemata. The minimally uniparametrically congruent butterfly schema features are a I–Dep…Ret–I and a 3:1 harmonic-rhythm ratio. The maximal uniparametrically congruent features are a I–Dep…V–I and a 1:1 harmonic-rhythm ratio. The textural grouping is a nonvariable congruent feature, formed at the level of regular functional harmonic change and at two immediately higher levels.

This chapter assesses the claim that butterfly schemata are parcels of features that are particular manifestations of the tendency for congruence in the Classical grammar. Since grammars have a spectrum of features of varying congruence that tend towards congruence (quantified through the rule of MC), the prevalence of multiparametrically congruent butterfly schemata in the Classical period suggests that the underlying cause of these schemata is the tendency for congruence. The survey of butterfly schemata during the period *c*. 1750–*c*. 1850 shows that butterfly schemata are almost three times more common in the Classical period (*c*. 1750–*c*. 1800) than the Romantic period (*c*. 1800–*c*. 1850). This suggests that the tendency for congruence causes butterfly schemata. An independent t-test showed that the findings were highly unlikely to have occurred due to chance (see Appendix C).

The second component of the survey examines the types of features found in butterfly schemata. It suggested that the type of features formed are also a product of the tendency for congruence. It was claimed that butterfly schemata of the Classical period should have more maximally than minimally congruent features (Section 1.4). Approximately 75% of butterfly schemata were maximally multiparametrically congruent. The remaining 25% of butterfly schemata had either minimally uniparametrically congruent features or an admixture of minimally and maximally uniparametrically congruent features. This confirms that maximally uniparametrically congruent features are more common than minimally uniparametrically congruent features in butterfly schemata of the Classical sample, providing further support that the tendency for congruence constrains the formation of butterfly schemata.

The top-down HS model shows that the features of the Classical grammar and butterfly schema are also formed due to cultural selection. While the universal tendency for congruence strongly influences features that form, culture also selects those features. Selection occurs in a particular top-down order, from whole to part. Features at the top of the hierarchy are more abstract than lower-level features. The rule of MC is manifested in universal rhythmic and pitch constraints at the top of the hierarchy, governing each feature in every level of the hierarchy. At the next level down, the particular type of metrical structuring (at the level of regular functional harmonic change and at two immediately higher levels) and diatonic space are selected. Following this, Classical chord progressions, textural grouping, and harmonic-rhythm ratios are selected. The butterfly schema is selected at the next level down, at the local schema level. The voice-leading schemata of Gjerdingen (1988, 2007) are dependent on butterfly schemata, being selected at the bottom of the hierarchy.

Chapter 6: Conclusions

The tendency for congruence permits a synthesis of associative-statistical and rationalist generative theories. The rule of MC was introduced as a quantification of the tendency for congruence in particular contexts. The butterfly schema and Classical grammar is a product of the tendency for congruence manifest in the particular form defined by the rule of MC. That is, butterfly schemata are formed from abstract features of the Classical grammar that are uniparametrically and multiparametrically congruent. The UC in features of the Classical grammar and butterfly schema was quantitatively measured. The multiparametric model of the butterfly schema incorporates the theory of tonal pitch space, develops the concept of textural grouping, and introduces the notion of harmonic-rhythm ratios. The rule of MC in the Classical grammar combines these uniparametrically congruent features into the butterfly schema, which are: chord progressions with short pitch-space distances from the tonic; textural grouping at the metrical level of regular functional harmonic change and at two immediately higher levels; simple harmonic-rhythm ratios formed at these levels; and congruent, but highly abstract, tonal and metrical structure.

A survey was conducted (reported in Section 5.2) which found that butterfly schemata are almost three times more common in the Classical-period sample than the Romantic-period sample. Since grammars are variably congruent but tend towards MC, the high popularity of multiparametrically congruent butterfly schemata (with maximally congruent features) in the Classical sample suggests that the tendency for congruence is a primary cause of these schemata. A top-down HS model of the butterfly schema (presented in Section 5.3) demonstrates that culturally conditioned features were cumulatively selected against each other, while being constrained by the tendency for congruence.

This chapter recapitulates the main conclusions and explores how the theories and models presented in this dissertation relate to corresponding studies and might influence future work. Section 6.1 shows how the tendency for congruence challenges associative-statistical approaches for modelling schemata, which have been argued to be limited explanatory theories of musical structure. Concrete schemata, such as the voice-leading schemata of Gjerdingen, form less consistent relationships with other features because they are not at the appropriate level of abstraction for consistent multiparametrically congruent interaction in the Classical grammar. Section 6.2 summarises generative theories in the light of the tendency for congruence. Generative theories, while showing the implicit congruence in musical structure, do not to provide a flexible framework for incorporating and combining variable forms of congruence and cultural conditioning. Section 6.3 assesses the findings and limitations of the survey of butterfly schemata, reviewing aspects of the methodology and its implications for future work. Section 6.4 evaluates the top-down HS model of the butterfly schema and discusses the limitations of a selective framework. Section 6.5 considers other grammars and schemata that can be analysed using the tendency for congruence, which might provide a more coherent system for future research.

6.1 The Tendency for Congruence and Associative-Statistical Theories of Schemata

A main argument developed through this dissertation is that congruence is not cognised through inductive inference but is rationally understood through an intuitive understanding of musical structure. The tendency for congruence is rather a product of cognition that influences grammars and schemata through history and in connection with culture, and so does not explain the mental act of music perception and cognition. Accordingly, associative-statistical and inductive theories do not provide explanations of musical cognition and knowledge. Associative-statistical models might show the statistical probability of an association network, but cannot explain the cause of such structures. However, in actuality, various inductive and statistical methods are often employed as mechanisms that propose to explain how musical structure is cognised. For example, inductive computational models (e.g., Cope (1991a, 1996, 2000b), Conklin and Witten (1995), and Temperley (2006)) were argued to be unsatisfactory approaches for explaining arrangements of uniparametrically and multiparametrically congruent musical structure (Section 3.1). Congruence in a single feature (UC) and between features (MC) must be intuitively cognised, but how this is done is not fully understood.

In Chapter 4 it was shown that the associative-statistical voice-leading schemata of Gjerdingen (1988, 2007) and Byros (2009a) are not located at the correct level of abstraction to form consistent multiparametrically congruent relationships with other features in the Classical grammar. While the schema theory of Gjerdingen (1988) is not presented as an inductive theory, it uses an empiricist methodology, positing that associative networks of features define schemata in certain times and places. Gjerdingen (2007) and Byros (2009a) also advance a theory of 'situated cognition', which suggests that listeners cognise music in particular ways depending on their cultural background. In Gjerdingen (1996), the notion of distinct 'courtly behaviors' is presented, where specific types of musical cognition are used. It is questionable whether these schema categories are really cognitive, rather than being merely 'situated', since a detailed cognitive theory is not presented in Gjerdingen (1988, 1996, 2007). However, Gjerdingen (1988) claims that schemata are categories formed in certain periods precisely because human brains, presumably through universal perceptual and cognitive mechanisms, interpret the distribution of schemata in similar ways (Gjerdingen, 1988, p. 99). This seems to be the opposite notion to situated cognition. The tendency for congruence permits the relationship between cognitive constraints and grammars and schemata to be more clearly and coherently defined, depending on the particular structure of grammars. The tendency for congruence is preferable approach to inductive theories because it can be applied to grammars as a general rule, before applying precise rules. It can incorporate the rules of particular grammars, while also considering universal constraints. The tendency for congruence offers a basis for the cognitive categorisation of schemata, without which features can only have a loose associative relationship and no explanatory framework to connect them.

Many of the voice-leading schemata of Gjerdingen (1988) can be construed as child schemata of the parent butterfly schemata. However, there are marked differences in the methodology, definition, and analysis of schemata in Gjerdingen (1988). In Gjerdingen (1988), noncongruent or partly congruent schemata are validated as prototypical voice-leading schemata (Sections 4.3.2 and 4.3.3 include discussions around this point). This is contrary to how schemata are defined using the rule of MC, which necessitates that these schemata are not butterfly schemata. Many voice-leading schemata which are congruent butterfly schemata in Gjerdingen (1988) are not validated as such due to the prioritisation of voice-leading structures. These noncongruent inclusions or congruent omissions of butterfly schemata in Gjerdingen (1988) show the fundamental difference between an associative-statistical methodology and the present analysis, which incorporates the tendency for congruence. Broadly, in Gjerdingen (1988), voice-leading events are given privileged attention, with corresponding inattention to the multiparametrically congruent grammatical structure. This is demonstrated through a review of schema theory analyses in Section 4.3, and supported in the survey in Section 5.2, which shows that certain voice-leading schemata (when construed as butterfly schemata) are not well represented in the Classical sample.

The lack of attention to the congruent grammatical structure in Gjerdingen (1988) means that the ontological status of voice-leading schemata is left unclear. It is evident that there must be generic and universal structuring of voice-leading schemata, otherwise schemata would not be understandable to listeners *a priori*. Similar butterfly schemata emerge in the grammars of other periods, such as in 'light' music of the nineteenth-century and, for example, popular music of WWII (which appears to be replete with butterfly schemata). This is because the tendency for congruence in those grammars regenerates similar structures. While it is feasible that schemata are *recognised* through perceptual-cognitive abstraction and association (as argued in Gjerdingen (1988)), it is not possible to fully ascertain how schemata are *generated* in the real world through associative methods. It is not established whether voice-leading schemata are nominal or ideal, abstract or concrete, or universal or particular, since an explanation in terms of their cognitive and cultural causes is not provided. From the results of the present survey (that voice-leading schemata are not common in the Classical period — when formalised as butterfly schemata) it is questionable whether voice-leading features are significant, either in cognition or in musical structure.

6.2 The Rule of MC and Generative Theories of Music

While generative theories generally unveil the implicit congruent structure of music, many rules therein are not generalisable to particular tonal grammars since they incorporate universal and general laws that can result in incoherent readings of those grammars. Congruence is not given sufficient focus in generative theories, since the aim of generative theories is to explain cognition. In generative theories, such as *GTTM*, a preference for congruence (incorporated through well-formedness and preference rules) is construed as part of the computational apparatus of cognition, rather than an indirect product of cognition, as is assumed in the present use of the tendency for congruence. Therefore, while generative theories tacitly incorporate the tendency for congruence, they often express relationships in structure that do not coherently reflect the flexibility of congruent structure found in various grammars and schemata. The tendency for congruence cannot be modelled through a finite set of algorithmic rules, since features are infinitely diverse and can combine in diverse combinations. Thus an emphasis on a non-inductive framework of the tendency for congruence is necessary to show how specific grammars are organised.

While generative models might demonstrate the tendency for congruence in hierarchical structure, they often do not fully account for cultural variation, or they concentrate on a single feature (as do many computational models). Thus the influence of culture is generalised and the interaction between features obscured. This contrasts with the present approach, which incorporates the tendency for congruence as a universal constraint, but examining particular cultures and focusing on particular multiparametric interaction. In computational musicology, generative approaches are incorporated into inductive frameworks to abstract grammatical rules from musical structure. For example, Longuet-Higgins and Steedman (1971), Krumhansl and Kessler (1982), and Maxwell (1992) use pattern-matching algorithms to establish implicit and stock tonal and harmonic structures. Also, the metrical algorithms of Povel and Essens (1985) and Rosenthal (1992) check passages for idealised metrical structures. While these harmonic and metrical computational models conduct automatic analyses that show the implicit tonal and metrical structure, they are limited due to their inductive frameworks, which cannot account for particular and variable multiparametrically congruent interaction.

Indeed, compuntational models that are uniparametric naturally do not permit insight into the multiparametric causes of musical structure. The biparametric grouping model of Tenney and Polansky (1980), the multiparametric implementation of *GTTM* in Hamanaka et al. (2005, 2007), and the multiparametric implementation of Schenkerian analysis and *GTTM* in Marsden (2001, 2007, 2010) are potentially more coherent models than uniparametric models because they examine the multiparametric interaction between features. However, cognition intuitively interprets the implicit multiparametrically congruent and noncongruent interaction in structure, which is a challenge to model computationally. Computational analyses can only show multiparametrically congruent relationships which are expressions of the tendency for congruence. They are often inflexible models that do not provide an explanation of how cognition intuitively understands multiparametrically congruent and noncongruent relationships.

Cooper and Meyer (1960) present a largely biparametric theory that examines the congruent and noncongruent interaction between rhythmic and metrical structure. *GTTM* improves on this framework with a multiparametric conception. However, as discussed, congruence is often not properly incorporated because it is presented through inflexible rules that do not show the variability of congruent structuring in grammars. Simply put, generative models present rules that are not generalisable. Peel and Slawson (1984, p. 291) argue that *GTTM* does not account for contrapuntal music, since the idiosyncratic and irregular grouping structure of contrapuntal music means that melodic phrase grouping is not a reliable cue for metrical structure. *TPS* and *CBMS* provide theories of contrapuntal grouping that are also unsatisfactory. In *TPS* and *CBMS*, contrapuntal grouping is not a feasible concept to account for the various types of contrapuntal structure that form, since their definition of the interaction of voices in texture presumes regular, duple hierarchical structure in the metrical hierarchy. In contrapuntal music, often it is only the beat level that yields a regular, duple hierarchical structure. Indeed, textural grouping and metrical structure in contrapuntal grammars of the Baroque period are grouped at low levels and is irregular at high levels. By contrast, in the Classical grammar and butterfly schema, which are either homorhythmic or heterorhythmic, textural grouping occurs at the level of regular functional harmonic change and at two immediately higher levels.

Well-formed tonal and metrical structure, while not considered to be a universal in *GTTM*, is thought to be a ubiquitous feature of tonal music (Lerdahl and Jackendoff, 1983, pp. 345–352). This is countered in the present dissertation since metrical well-formedness — regular, hierarchical, and grid-like structure — is shown to be a universal constraint that is variably instantiated. Metrical structure is highly dependent on texture, which can be grouped in various particular ways, meaning metrical structure can likewise be grouped in distinct ways. While metrical structure is also a product of chord progression and harmonic rhythm, textural grouping is perhaps its strongest constraint. The present view of metrical structure challenges the essentialist notion of metre in *GTTM*, *CBMS*, and London (2004), and many computational approaches, which assert that it is necessarily a well-formed structure. Likewise, tonal structure is also highly dependent on context. Universally constrained, but also culturally situated basic spaces, constrain the pitch structure of grammars, and are used to model chord progression in the butterfly schema.

Generative models, such as *GTTM* and *CBMS*, incorporate Gestalt principles (Wertheimer, 1923; Koffka, 1935) and auditory stream segregation (Bregman, 1990) as fundamental principles of musical cognition, necessary to parse grouping and metrical structure. Gestalt principles and auditory stream segmentation, while important universal constraints on the perception and cognition of events, are not comprehensive explanations of grouping structure (see Section 3.2.1). Textural grouping can be highly differentiated depending on the grammatical context, due to its interaction with culture. Viewing melodic or textural grouping solely through the Gestalt principles of similarity, proximity, and symmetry limits a reading of the variety of structuring that form in various grammars. In the congruent model of the butterfly schema, the Gestalt principles of proximity and similarity are assumed to be in some way necessary to govern how texture is grouped, but culture determines the level at which textural grouping and metrical grid-like structure takes place.

A further criticism of generative models, but which also applies to schema theories, is that the sample to which the theories apply are not exactly delimited (a necessary stipulation for statistical analyses of style (Meyer, 1989, p. 57).[39] In *GTTM*, the sample is not made historically or geographically explicit, only vaguely insinuated by the designation of 'tonal', and suggested with reference to the 'common practice period', *c*. 1600–1900. Narrower limits are necessary because in an unlimited data field the results cannot be objective, and data might be cherry-picked to fit the hypothesis. Defining the sample as 'tonal' confounds and obscures the body of music to which the theory is directed. The designation of the common practice period is also problematic because the rules of *GTTM* are most applicable to eighteenth-century music (Peel and Slawson, 1984, p. 291). Some of the rules apply to certain grammars of Western music while others do not. A theory of tonal music must account for the vast array of grammars and their features, such as the modal harmonic structure of the Renaissance period, or neo-Riemannian systems of nineteenth-century music,[40] as well as their various types of textural and metrical profiles. However, there is perhaps no way of incorporating the divergent types of structuring into a single generic generative theory, since congruence is variably manifest. While *TPS* and *CBMS* are more sophisticated than *GTTM* in this respect, incorporating Western and non-Western styles outside the common practice period, they present general and universal rules that do not apply to all grammars. As mentioned, diverse grammatical features can be more coherently modelled through the tendency for congruence, which can be flexibly applied to particular contexts through the rule of MC.

While *GTTM* incorporates harmonic change in its system of rules, it does so under the broad rubric of a change of accentuation (Lerdahl and Jackendoff, 1983, pp. 76–89). *GTTM*, *TPS*, and *CBMS* do not incorporate harmonic rhythm as a specific feature of grammars, yet it is posited to be an important constraint on musical structure by many music theorists, such as Lester (1986), Rothstein (1989), Swain (2002), McKee (2004), and Mirka (2009). Regular harmonic rhythm and harmonic-rhythm ratios have been demonstrated to be

[39] The chapter 'Statistics' in Meyer (1989, pp. 57–65) explores issues concerning the demarcation of samples.

[40] Lerdahl (2001) widens the scope of *GTTM* to include chromatic relationships (incorporating neo-Riemannian theory), as well as presenting hypotheses about how grammatical theories could be applied to music of other periods and cultures.

uniparametrically congruent features necessary for MC in the Classical grammar and butterfly schema (Section 5.1.3). These are formed at the level of regular functional harmonic change and at two immediately higher levels. Particular harmonic-rhythm ratios (1:1 and 3:1) are more significant features in the Classical grammar than in the Baroque and Romantic grammars, because in the grammars of these periods harmonic rhythm is less consistently aligned with metrical structure. In many Renaissance contrapuntal grammars, harmonic rhythm is often not a primary feature of grammar but a consequence of voice-leading interaction (Berry, 1976, p. 200; Swain, 2002, pp. 129–141). That harmonic rhythm is dependent on culture points to the necessity of incorporating the tendency for congruence rather than using systematised cognitive rules, which can be incoherent.

Notwithstanding its application in the butterfly schema and Classical grammar, tonic harmony prolongation at middleground and global levels is not a universal of all grammars (although significant for the Classical grammar; see rule i, Section 5.1.1.2). In *GTTM*, the specific rules of prolongational reduction, time-span reduction, and the interaction principle determine the pitch and rhythmic structure of the butterfly schema and Classical grammar. However, these distinct rules, which show the reciprocal interaction between pitch and rhythmic stability at certain levels, are not generalisable. Grammars can either prioritise pitch prolongation or rhythmic structure (as argued in Yeston (1976, pp. 4–5). *GTTM* does not flexibly account for various dominances between rhythm and pitch. This is again a product of the wider approach of generative theories that provide generalisable cognitive rules at the expense of showing the variability of grammatical structure. By contrast, the tendency for congruence merely shows the influence of cognition on musical structure, permitting more varied and coherent interactions between the domains of rhythm and pitch. However, despite Peel's and Slawson's (1984, p. 275) and Harvey's (1985, p. 295) claims that the time-span reduction and prolongational components of *GTTM* are generally not as revealing as the grouping and metrical components, the insights on the congruent and noncongruent interaction between time-span reduction and prolongational reduction are the perhaps the most interesting contributions. These are useful to define the interaction between pitch and rhythmic stability in the regions of the chord progression of the butterfly schema.

The present portrayal of real-time listening (Section 4.1.2) is generally coherent with generative theories in the sense that perception and cognition generates larger, abstract global structures from simpler, less abstract congruent features. The present view of perception and cognition also follows the wider cognitive science tradition that musical structure is a type of internal language (I-language), rather than external language (E-language) of communication (explained in Chomsky (2006, p. 175) for language). However, a theoretical framework for how this is done is not provided through the tendency for congruence. Also, musical structure is fundamentally different to generative theories of language because music involves the interaction of many constituent hierarchies, such as chord progression and metrical structure, which can interact in infinite ways, and so is probably not characterised by a single well-formed structure.

The tendency for congruence has no explicit cognitive rules, and no specific well-formed structures. The rule of MC is an expression of the tendency for congruence in specific contexts. MC is variously manifest in grammars, and so the rule of MC can be flexibly applied to particular contexts. The rule of MC is incorporated as part of a non-inductive explication of the Classical grammar. The tendency for congruence and the rule of MC do not explain cognition, but rather show the influence of cognition on musical structure. The implicit congruent structure in and between features is intuitively cognised. However, it is not possible to fully explain how this is achieved. Since musical structure is infinite, understanding must be rational and deductive. In future research in musical grammars, the present use of the non-inductive tendency for congruence (and the rule of MC) might provide a useful approach for coherently representing musical structure.

6.3 The Survey of Butterfly Schemata

This dissertation set out the claim that butterfly schemata emerge more frequently in the Classical samples of European instrumental works than in the Romantic samples. As reported in Section 5.2, butterfly schemata do indeed emerge in the Classical samples (22% of movements) more frequently than in the Romantic samples (7.6% of movements) (the full survey is presented in Appendix A). This suggests that the particular manifestation of the tendency for congruence produces more butterfly schemata in the Classical period than in the Romantic period. The influence of the tendency for congruence was also supported by the type of butterfly schemata found in the Classical samples. Maximally congruent butterfly schemata were more commonly found than minimally congruent schemata, again suggesting that the tendency for congruence is their underlying cause.

Schemata can only be called such if they are common entities of a particular time and place. From the survey, the Meyer (1–7…4–3) voice-leading schema, when modelled as a grammatical child schema of the parent butterfly schema, was found in fewer than a quarter of all butterfly schemata in the Classical samples. The survey found that other voice-leading schemata of Gjerdingen (2007), the Sol–Fa–Mi, Aprile, Jupiter, Pastorella, and Do–Re…Re–Me, when defined as child schemata of the grammatically congruent butterfly schemata, are also sparsely represented in the Classical-period samples. Considering the diversity of voice-leading structures conceivable in the chord progression of the butterfly schema, this is a better than chance representation, yet the underlying implicit grammatical structure, which has been shown to govern the formation of the butterfly schema, provides a more convincing explanatory framework. Network relationships of voice-leading schemata are not revelatory about the cause of the interaction of features (as suggested in Cohn and Dempster (1992)). Associative-statistical theories of voice-leading schemata might explain the emergent voice-leading structure, but do not examine the root cause of the ubiquitous grammatical structure, which is here proposed to be the tendency for congruence. It is possible that in the pre-Classical period voice-leading features might have had greater significance, as suggested by the predominance of voice-leading rules during this time, such as the *Règle de l'octave* (Gjerdingen, 2007, pp. 467–470; Sanguinetti, 2012, pp. 113–125). However, these are presumably not as important in the Classical instrumental grammar.

Butterfly schemata were also found in the Romantic-period samples. Although the output of butterfly schemata in the Classical grammar (22%) was much higher than that of the Romantic grammar (7.6%), the small quantity of butterfly schemata found in the Romantic grammar can be explained as vestiges of the Classical grammar that are residual in the latter grammar. During the course of the Romantic period, the Classical-period features are gradually usurped by emergent congruent features of the Romantic grammar. While these grammars are fuzzy at the periodic boundaries, there is a marked difference in the output of butterfly schemata of each period. An independent t-test established that the means of the samples were significantly different, which suggests that these results were unlikely to have occurred by chance (see Appendix C). The statistical test was an appropriate technique to establish the reliability of the means of the samples of each period, and could be adopted in future corpus studies that compare musical grammars and schemata.

The grammatical evolution between the Classical and Romantic periods is epitomised in the output of Beethoven, who was active in both periods. Beethoven's style undergoes a radical shift around the year 1804. From this time, butterfly schemata, which had formed a significant portion of his eighteenth-century samples, radically decline in popularity (as can be seen in his Romantic-period samples in Appendix B). While this difference corresponds with various accounts of Beethoven's individual stylistic evolution between these periods (e.g., Rosen (1971)), his style change is also explicable as part of the quantitative evolutionary change in the grammar in general during this time. This change is brought about through the gradual cultural selection of new congruent features in the Romantic instrumental grammar.

Butterfly schemata form only a small part of the structure of individual works, which form a modest percentage of the Classical-period sample as a whole. It may therefore be questionable whether the butterfly schema is a significant representation of grammatical structure during this period. However, the butterfly schema is not an independent structure but contains features that are ubiquitous in the whole Classical grammar, formed through the tendency for congruence and cultural selection. Butterfly schemata are significant in terms of the arrangement of their features within the grammatical system. The tendency for congruence constrains the form and interaction of every feature of the grammar and schemata. While two features of the Classical grammar, such as textural grouping and harmonic rhythm, might frequently interact biparametrically congruently, producing bi-parametric schemata, the requirement that all three abstract congruent features occur together (chord progression, textural grouping, and harmonic-rhythm ratio), in the context of the highly abstract congruent tonal and metrical structure (necessary to satisfy the rule of MC), significantly lowers the probability of its emergence. Although the popularity of a schema is directly related to the degree of congruence in and between its features, increasing the quantity of features that define a schema significantly lowers its popularity. This framework usefully relates schemata to grammatical structure, showing how they are intrinsically connected.

6.4 The Top-Down Hierarchical Selection Model of the Butterfly Schema

The model of top-down hierarchical selection (HS) shows how features exist in the Classical grammar and butterfly schema, revealing the order of dependency (selection) of features. The top-down HS model illustrates how features at various levels of the Classical grammar and butterfly schema are arranged in a particular order, where each feature is cumulatively selected in a top down hierarchy. In top-down HS, the order of selection is dictated by the action of culture. This contrasts with the perception and cognition of these structures in real-time listening, where there is bottom-up causation.

The top-down HS mirrors biological evolution: as universal physical laws constrain natural selection, the universal cognitive tendency for congruence constrains top-down HS. Top-down HS shows how the cognitive tendency for congruence can be incorporated into a cultural selection model. The rule of MC, which is a formal expression of this tendency, constrains all levels of selection. While governed by the rule of MC, tonal structure and metrical structure are culturally selected, determining the formation of less abstract features lower in the hierarchy. The rule of MC and top-down HS do not directly counter the notion of compositional creativity, but nonetheless are the main causes of musical structure. Corresponding to the theory of 'selection by consequences' (Skinner, 1981), in the past, low-level features of the hierarchy were shaped and maintained by high-level features. Thus, the features at low levels do not exist as direct responses to high-level features (by composers), or as responses to those features in particular pieces, but are the consequence of a broader cultural selection history.

The top-down HS model shows the formation and interaction of features of the Classical grammar and butterfly schema that cannot be understood by the tendency for congruence alone. It elucidates how features are culturally conditioned, which is the top-down cause of grammars and schemata. For example, the metrical structure of the Classical period (regular duple hierarchy at the level of regular functional harmonic change and at two immediately higher levels) is constrained by tendency for congruence, but ultimately selected by the top-down action of culture. In real-time listening, the ordering of structure is the converse, which is the bottom-up cause of metrical and tonal structure. Real-time listening involves the bottom-up perception and cognition of abstract features — chord progression, harmonic rhythm, and textural grouping — from which the highly abstract tonal and metrical structure are conceived (Section 4.1.2 explains bottom-up real-time listening). This ordering in real-time listening must be universal because it explains how complex and variable musical structure is understood. However, musical structure is ultimately caused by top-down cultural selection, which governs the form and interaction of lower-level features in the hierarchy, and can differ between grammars.

Top-down HS and the rule of MC cannot account for all types of structuring that emerge in music. While these are fundamental constraints on structure it is possible for arbitrary and noncongruent structuring to form in Classical instrumental music that is contrary to the formal models, but usually at lower levels, which are less stable than higher levels. For example, noncongruent local schemata, such as the Neapolitan sixth chord preceding a cadence, might frequently occur at lower levels, although such noncongruent structures are infrequent at higher levels. Generally, however, in the Classical grammar, noncongruent structures are infrequent at all levels. An important axiom of the top-down HS model is that emergence does not routinely occur in grammars, since the order of selection is normative. Noncongruent top-down HS does not generally occur in the Classical grammar and butterfly schema, since the features that are selected in the hierarchy follow the rule of MC. Notwithstanding, it is theoretically feasible that selection occurs for features that are noncongruent in other grammars, contradicting the tendency for congruence, providing an interesting confliction that would require a noncongruent selection hierarchy. Artificial grammars, such as those that use serial techniques, might present such cases, where multiparametrically *noncongruent* features are selected for (a similar argument is made in Lerdahl (1988a)).

The analyses presented in Gjerdingen (1988, 2007) contend that voice-leading features are significant for schemata. However, Gjerdingen's voice-leading analyses are shown to be, in some way, incoherent with the present explanations of the multiparametrically congruent grammatical structure (in Sections 4.3 and 5.2). Indeed, voice-leading schemata, formulated as multiparametrically congruent butterfly schemata, are the least significant culturally conditioned features, occupying the lowest position in the top-down HS model. It is questionable whether they should occupy a place in general because the other multiparametrically congruent features of the butterfly schema and Classical grammar infrequently select them.

6.5 The Tendency for Congruence in Other Grammars and Schemata

Chapters 4 and 5 comprise comparative analyses that illustrate the difference in multiparametrically congruent structure between the Classical, Romantic, and Baroque grammars, as well as comparisons with voice-leading interpretations of schemata. Since the tendency for congruence is an implicit abstract tendency of structure, it presents an elusive constraint requiring further investigation to uncover a wider significance. A more focused examination of the contrapuntal grammars of the Baroque and Renaissance could show more specifically how the tendency for congruence is manifest. Renaissance grammars, such as the sacred compositions of Italian composers (e.g., Giovanni Pierluigi de Palestrina (*c.* 1525–1594)) have textural structures that are ostensibly highly noncongruent at high levels of texture but which interact congruently with the modal structure at low levels. In the investigation of grammars and schemata, it is possible that new explications of multiparametrically congruent relationships might be uncovered. Harmonic rhythm is a feature that was not previously incorporated in associative-statistical or generative theories, but which has been shown to be an essential feature in the present examination of the

Classical grammar and butterfly schema. Novel features can be expounded through using the present approach in other grammars and schemata.

Many grammars and their corresponding schemata are similar to the Classical grammar and butterfly schema. 'Light' music and opera of the nineteenth century and popular Western music of the twentieth century often incorporate butterfly schemata, including the music of composers such as Vincenzo Bellini (1801–1835) or Guiseppe Verdi (1813–1901) (although art music of the nineteenth century generally comprises many different types of grammars and schemata), but a statistical analysis is required to provide evidence of a correlation. Popular songs of WWII ostensibly contain butterfly schemata, and also use other voice-leading schemata, similar to those expounded in Gjerdingen (2007). Such schemata require further analysis in order to find whether the apparent similarity continues throughout corpora, providing evidence of grammatical homogeneity and also the tendency for congruence. An examination of non-Western grammars might help to determine how the tendency for congruence might be manifest, providing an alternative to generative theories of music cognition. This would provide an expansion of understanding of the relationship between musical grammars and schemata, building on, but also challenging, associative-statistical and generative theories of schemata.

∞

This thesis has resisted the notion that the tendency for congruence might form part of an inductive or context-free theory of musical understanding, even though a rational appreciation of congruence and noncongruence is an important assumption about how music is cognised. Congruence and noncongruence are *a priori* concepts that must be intuitively interpreted by human rationality because they can be understood even within limitless variation. However, it is unlikely that there can be a systematic theory that shows how congruence and noncongruence is understood in every circumstance.

A novel theoretical framework for explaining the combination of the constraints of cognition and culture on musical structure has been presented. The tendency for congruence, the multiparametrically congruent model of the butterfly schema, and the top-down HS model provide a way to unite the conceptions of generative and associative-statistical theories into a richer framework for examining musical structure. The limitations of generative theories (which are often too generalised and focussed on providing an explanation of cognition) and the limitations of associative-statistical theories (that eschew the influence of universal generative capacities) are overcome through the incorporation of the tendency for congruence. The tendency for congruence and top-down HS are generalisable, applying broad principles of cognition that also admit the integration of highly specialised cultural features. The theories and methods used in this thesis can be applied to further understanding of the structure of other grammars and schemata, where the aim is to develop cognitively influenced, holistic systems of interacting features.

Appendix A

Results of the Statistical Survey of Butterfly Schemata in European Instrumental Music (*c.* 1750–*c.* 1800)

The following is a list of ascriptions used in the statistical analysis of butterfly schemata:

- Butterfly = a butterfly schema found
- Nil = no butterfly schemata found
- Max. chord progression = maximally uniparametrically congruent chord progression (T–Dep…V–T)
- Min. chord progression = minimally uniparametrically congruent chord progression (T–Dep…Ret–T)
- 1:1 = maximally uniparametrically congruent harmonic-rhythm ratio
- 3:1 = minimally uniparametrically congruent harmonic-rhythm ratio
- Meyer = voice-leading schema
- Aprile = voice-leading schema
- Pastorella = voice-leading schema
- i, ii, iii, iv, etc. = movement number

A. 1 Haydn String Quartets

B♭ major, Hob.III:1 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: Nil, iv (Menuetto): butterfly, 1:1, min. chord progression, bars 1–4, IV (Trio): butterfly, 1:1, min. chord progression, bars 1–4, v: Nil.

E♭ major, Hob.III:2 (5 movs.) (*c.* 1757–62) – i: butterfly, 1:1, max. chord progression, bars 1–4, ii (Menuetto): butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil, v: Nil.

D major, Hob.III:3 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: Nil, iv (Menuetto): butterfly, 1:1, min. chord progression, bars 1–4, v: Nil.

G major, Hob.III:4 (5 movs.) (*c.* 1757–62) – i: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, ii: Nil, iii: Nil, iv: Nil, v: Nil.

B♭ major, Hob.III:5 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: Nil, iv: Nil, v: Nil.

C major, Hob.III:6 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: Nil, iv: Nil, v: Nil.

A major, Hob.III:7 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, min. chord progression, bars 1–4, v: Nil.

E major, Hob.III:8 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: Nil, iv: Nil, v: Nil.

E♭ major, Hob.III:9 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: butterfly (Meyer), 1:1, max. chord progression, bars 1–2, iv: Nil, v: butterfly, 1:1, max. chord progression, bars 1–2.

F major, Hob.III:10 (5 movs.) (*c.* 1757–62) – i: Nil, ii: Nil, iii: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, iv (Menuetto): butterfly, 1:1, max. chord progression, bars 1–4, v: Nil.

D major, Hob.III:11 (5 movs.) (*c.* 1757–62) – i: butterfly, 1:1, max. chord progression, bars 92–95, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–4, v: Nil.

B♭ major, Hob.III:12 (5 movs.) (*c.* 1757–62) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 13–16, iii: Nil, iv: Nil, v: butterfly, 1:1, max. chord progression, bars 8–11.

E major, Hob.III:13 (4 movs.) (1767) – i: Nil, ii (Trio): butterfly, 1:1, max. chord progression, bars 29–32, iii: Nil, iv: Nil.

C major, Hob.III:14 (3 movs.) (1767) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 229–236.

G major, Hob.III:15 (4 movs.) (1767) – i: Nil, ii: Nil, iiio: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–4.

B♭ major, Hob.III:16 (4 movs.) (1767) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

F major, Hob.III:17 (4 movs.) (1767) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A major, Hob.III:18 (4 movs.) (1767) – i: Nil, ii: Nil, iii (Menuetto): butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

C major, Hob.III:19 (4 movs.) (1769) – i: Nil, ii (Menuetto): butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

E♭ major, Hob.III:20 (4 movs.) (1769) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

G major, Hob.III:21 (4 movs.) (1769) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–4.

D minor, Hob.III:22 (4 movs.) (1769) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B♭ major, Hob.III:23 (4 movs.) (1769) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A major, Hob.III:24 (4 movs.) (1769) – i: Nil, ii: (Menuetto), butterfly, 1:1, max. chord progression, bars 9–12, iii: Nil, iv: Nil.

E major, Hob.III:25 (4 movs.) (1771) – i: Nil, ii (Trio): butterfly, 1:1, max. chord progression, bars 51–54, iii: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, iv: Nil.

F major, Hob.III:26 (4 movs.) (1771) – i: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: butterfly (Aprile), 1:1, max. chord progression, bars 1–4.

E♭ major, Hob.III:27 (4 movs.) (1771) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

C minor, Hob.III:28 (4 movs.) (1771) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

G major, Hob.III:29 (4 movs.) (1771) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D major, Hob.III:30 (4 movs.) (1771) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

E♭ major, Hob.III:31 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

C major, Hob.III:32 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

G minor, Hob.III:33 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D major, Hob.III:34 (4 movs.) (1772) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 21-24, iv: Nil.

F minor, Hob.III:35 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A major, Hob.III:36 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B minor, Hob.III:37 (4 movs.) (1781) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

E♭ major, Hob.III:38 (4 movs.) (1781) – i: Nil, ii: Nil, iii (Trio): butterfly (Meyer), 1:1, max. chord progression, bars 35–38, iv: Nil.

C major, Hob.III:39 (4 movs.) (1781) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B♭ major, Hob.III:40 (4 movs.) (1781) – i: Nil, ii: Nil, iii: Nil, iv: butterfly (Aprile), 1:1, max. chord progression, bars 37–40.

G major, Hob.III:41 (4 movs.) (1781) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

D major, Hob.III:42 (4 movs.) (1781) – i: Nil, ii: butterfly (Aprile), 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

D major, Hob.III:43 (4 movs.) (1785) – i: butterfly, 1:1, max. chord progression, bars 40–43, ii: Nil, iii: Nil, iv: Nil.

B♭ major, Hob.III:44 (4 movs.) (1787) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

C major, Hob.III:45 (4 movs.) (1787) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

E♭ major, Hob.III:46 (4 movs.) (1787) – i: butterfly (Do–Re…Re–Me), 1:1, max. chord progression, bars 1–2, ii: Nil, iii: Nil, iv: Nil.

F♯ minor, Hob.III:47 (4 movs.) (1787) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

F major, Hob.III:48 (4 movs.) (1787) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D major, Hob.III:49 (4 movs.) (1787) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

B♭ major, Hob.III:50 (1787) – i: butterfly (Meyer), 1:1, max. chord progression, bars 1–2

C minor, Hob.III:51 (1787) – i: Nil.

E major, Hob.III:52 (1787) – i: butterfly, 1:1, max. chord progression, bars 3–6.

F minor, Hob.III:53 (1787) – i: butterfly (Pastorella?), 1:1, max. chord progression, bars 1–4.

A major, Hob.III:54 (1787) – i: Nil.

G minor, Hob.III:55 (1787) – i: Nil.

E♭ major, Hob.III:56 (1787) – i: Nil.

C major, Hob.III:57 (4 movs.) (1788?) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

G major, Hob.III:58 (4 movs.) (1788?) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

E major, Hob.III:59 (4 movs.) (1788?) – i: Nil, ii: Nil, iii: butterfly, 3:1, min. chord progression, bars 1–8, iv: butterfly, 1:1, max. chord progression, bars 1–4.

A major, Hob.III:60 (4 movs.) (1788) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

F minor, Hob.III:61 (4 movs.) (1788) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B♭ major, Hob.III:62 (4 movs.) (1788) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D major, Hob.III:63 (4 movs.) (1790) – i: butterfly, 1:1, max. chord progression, bars 9–16, ii: Nil, iii: Nil, iv: Nil.

E♭ major, Hob.III:64 (4 movs.) (1791) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 37–44, iv: Nil.

C major, Hob.III:65 (4 movs.) (1790) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

G major, Hob.III:66 (4 movs.) (1790) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B♭ major, Hob.III:67 (4 movs.) (1790) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B minor, Hob.III:68 (4 movs) (1790) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

B♭ major, Hob.III:69 (4 movs.) (1793) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D major, Hob.III:70 (4 movs.) (1793) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil

E♭ major, Hob.III:71 (4 movs.) (1793) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

C major, Hob.III:72 (4 movs.) (1793) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–8, iii: Nil, iv: Nil.

F major, Hob.III:73 (4 movs.) (1793) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

G minor, Hob.III:74 (4 movs.) (1793) – i: butterfly, 1:1, max. chord progression, bars 169–172, ii: Nil, iii: Nil, iv: Nil.

G major, Hob.III:75 (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D minor, Hob.III:76 (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 2–5, iv: Nil.

C major, Hob.III:77 (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

B♭ major, Hob.III.78 (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

D major, Hob.III.79 (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

E♭ major, Hob.III:80 (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–4.

G major, Hob.III:81 (4 movs.) (1799) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

F major, Hob.III:82 (4 movs.) (1799) – i, Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

D minor, Hob.III:83 (2 movs.) (1803) – i: Nil, ii: Nil.

A.2 Haydn Piano Sonatas

No. 1 in G major, XVI:8 (4 movs.) (1766) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 2 in C major, XVI:7 (3 movs.) (1766) – i: Nil, ii: Nil, iii: Nil.

No. 3 in F major, XVI:9 (3 movs.) (1766) – i: Nil, ii: Nil, iii: Nil.

No. 5 in G major, XVI:11 (3 movs.) (1767) – i: butterfly, 1:1, max. chord progression, bars 25–28, ii: Nil, iii: Nil.

No. 6 in C major, XVI:10 (3 movs.) (1767) – i: Nil, ii: Nil, iii: Nil.

No. 8 in A major, XVI:5 (3 movs.) (1750–55) – i: butterfly (Meyer), 1:1, max. chord progression, bars 38–41, ii: Nil, iii: Nil.

No. 9 in D major, XVI:4 (2 movs.) (1765) – i: Nil, ii: Nil.

No. 10 in C major, XVI:1 (3 movs.) (1750–55) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 21–24.

No. 11 in B♭ major, XVI:2 (3 movs.) (1760) – i: Nil, ii: Nil, iii: Nil.

No. 12 in A major, XVI:12 (3 movs.) (1767) – i: butterfly, 1:1, max. chord progression, bars 1–4, ii: Nil, iii: Nil.

No. 13 in G major, XVI:6 (4 movs.) (1760–66) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 14 in C major, XVI:3 (3 movs.) (1765) – i: Nil, ii: Nil, iii: Nil.

No. 15 in E major, XVI:13 (3 movs.) (1760–67) – i: Nil, ii: Nil, iii: Nil.

No. 16 in D major, XVI:14 (3 movs.) (1760–67) – i: Nil, ii: Nil, iii: Nil.

No. 19 in F major, XVI:47bis (3 movs.) (1765–67) – i: Nil, ii: Nil, iii: Nil.

No. 20 in B♭ major, XVI:18 (2 movs.) (1771–73) – i: Nil, ii: Nil.

No. 29 in E♭ major, XVI:45 (3 movs.) (1766) – i: butterfly, 1:1, max. chord progression, bars 19–20, ii: Nil, iii: Nil.

No. 30 in D major, XVI:19 (3 movs.) (1767) – i: butterfly (Meyer), 1:1, max. chord progression, bars 4–5, ii: Nil, iii: Nil.

No. 31 in A♭ major, XVI:46 (3 movs.) (1767–70) – i: Nil, ii: Nil, iii: Nil.

No. 32 in G minor, XVI:44 (2 movs.) (1771–73) – i: Nil, ii: Nil.

No. 33 in C minor, XVI:20 (3 movs.) (1771) – i: Nil, ii: Nil, iii: Nil.

No. 34 in D major, XVI:33 (3 movs.) (1778) – i: Nil, ii: Nil, iii: Nil.

No. 35 in A♭ major, XVI:43 (3 movs.) (1783) – i: Nil, ii: Nil, iii: Nil.

No. 36 in C major, XVI:21 (3 movs.) (1773) – i: butterfly, 1:1, max. chord progression, bars 1–4, ii: Nil, iii: Nil.

No. 37 in E major, XVI:22 (3 movs.) (1773) – i: Nil, ii: Nil, iii: Nil.

No. 38 in F major, XVI:23 (3 movs.) (1773) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4.

No. 39 in D major, XVI:24 (3 movs.) (1773) – i: Nil, ii: Nil, iii: Nil.

No. 40 in E♭ major, XVI:25 (2 movs.) (1773) – i: butterfly, 1:1, max. chord progression, bb. 8–10, ii: Nil.

No. 41 in A major, XVI:26 (3 movs.) (1773) – i: Nil, ii: Nil, iii: Nil.

No. 42 in G major, XVI:27 (3 movs.) (1774–76) – i: butterfly, 1:1, min. chord progression, bars 1–4, ii: Nil, iii: Nil.

No. 43 in E♭ major, XVI:28 (3 movs.) (1774–76) – i: Nil, ii: butterfly, 1:1, min. chord progression, bb. 1–4, iii: Nil.

No. 44 in F major, XVI:29 (3 movs.) (1774) – i: Nil, ii: butterfly, 1:1, min. chord progression, bars 1–2, iii: Nil.

No. 45 in A major, XVI:30 (3 movs.) (1774–76) – i: Nil, ii: Nil, iii: Nil.

No. 46 in E major, XVI:31 (3 movs.) (1774–76) – i: butterfly, 1:1, max. chord progression, bars 5–6, ii: Nil, iii: Nil.

No. 47 in B minor, XVI:32 (3 movs.) (1776) – i: Nil, ii: Nil, iii: Nil.

No. 48 in C major, XVI:35 (3 movs.) (1780) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–2, iii: Nil.

No. 49 in C♯ minor, XVI:36 (3 movs.) (1772) – i: Nil, ii: Nil, iii: Nil.

No. 50 in D major, XVI:37 (3 movs.) (1780) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4.

No. 51 in E♭ major, XVI:38 (3 movs.) (1780) – i: Nil, ii: butterfly (Meyer), 1:1, max. chord progression, bars 1–2, iii: Nil.

No. 52 in G major, XVI:39 (3 movs.) (1780) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 8–11, iii: Nil.

No. 53 in E major, XVI:45 (3 movs.) (1784) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil.

No. 54 in G major, XVI:40 (2 movs.) (1784) – i: Nil, ii: Nil.

No. 55 in B♭ major, XVI:41 (2 movs.) (1784) – i: butterfly (Meyer), 1:1, max. chord progression, bars 8–12, ii: Nil.

No. 56 in D major, XVI:42 (2 movs.) (1784) – i: Nil, ii: Nil.

No. 57 in F major, XVI:47 (3 movs.) (1788) – i: butterfly, 1:1, max. chord progression, bars 9–12, ii: Nil, iii: Nil.

No. 58 in C major, XVI:48 (2 movs.) (1789) – i: Nil, ii: Nil.

No. 59 in E♭ major, XVI:49 (3 movs.) (1789) – i: butterfly (Meyer), 1:1, max. chord progression, bars 25–28, ii: Nil, iii: butterfly, 1:1, min. chord progression, bars 1–4.

No. 60 in C major, XVI:50 (3 movs.) (1794) – i: Nil, ii: Nil, iii: Nil.

No. 61 in D major, XVI:51 (2 movs.) (1794) – i: Nil, ii: Nil.

No. 62 in E♭ major, XVI:52 (3 movs.) (1794) – i: Nil, ii: Nil, iii: Nil.

A.3 Mozart Symphonies

No.1, E♭ major, K. 16 (3 movs.) (1764–65) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars1–4, iii: Nil.

No. 2, B♭ major, K. 17 (4 movs.) (1787) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–2, iii: Nil, iv: Nil.

No. 3, E♭ major, K. 18 (3 movs.) (1764) – i: Nil, ii: Nil, iii: Nil.

No. 4, D major, K. 19 (3 movs.) (1765) – i: Nil, ii: Nil, iii: Nil.

No. 5, B♭ major, K. 22 (3 movs.) (1765) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil.

No. 6, F major, K. 43 (4 movs.) (1767) – i: Nil, ii: butterfly (Aprile/Meyer), 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

No. 7, D major, K. 45 (4 movs.) (1768) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 8, D major, K. 48 (4 movs.) (1768) – i: Nil, ii: Nil, iii: butterfly, 1:1, min. chord progression, bars 1–4, iv: Nil.

No. 9, C major, K. 73 (4 movs.) (1769/72) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 10, G major, K. 74 (3 movs.) (1770) – i: Nil, ii: Nil, iii: Nil.

No. 11, D major, K. 84 (3 movs.) (1770) – i: Nil, ii: Nil, iii: Nil.

No. 12, G major, K. 110 (4 movs.) (1771) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 13, F major, K. 112 (4 movs.) (1771) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

No. 14, A major, K. 114 (4 movs.) (1771) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 9–12, iv: butterfly (Meyer), 1:1, max. chord progression, bars 74–80.

No. 15, G major, K. 124 (4 movs.) (1772) – i: butterfly (Do–Re…Re–Mi), 1:1, max. chord progression, bars 1–4, ii: butterfly (Meyer), 1:1, max. chord progression, bars 11–12, iii: Nil, iv: Nil.

No. 16, C major, K. 128 (3 movs.) (1772) – i: butterfly, 1:1, min. chord progression, bars 2–5, ii: Nil, iii: Nil.

No. 17, G major, K. 129 (3 movs.) (1772) – i: Nil, ii: Nil, iii: butterfly (Meyer), 1:1, max. chord progression, bars 46–49.

No. 18, F major, K. 130 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 19, E♭ major, K. 132 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 20, D major, K. 133 (4 movs.) (1772) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 21, A major, K. 134 (4 movs.) (1772) – i: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, ii: butterfly (Meyer), 1:1, min. chord progression, bars 1–4, iii: Nil, iv: Nil.

No. 22, C major, K. 162 (3 movs.) (1773) – i: Nil, ii: Nil, iii: Nil.

No. 23, D major, K. 181 (3 movs.) (1773) – i: Nil, ii: butterfly, 3:1, min. chord progression, bars 15–22, iii: Nil.

No. 24, B♭ major, K. 182 (3 movs.) (1773) – i: Nil, ii: Nil, iii: Nil.

No. 25, G minor, K. 183 (4 movs.) (1773) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii (Trio): butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

No. 26, A major, K. 184 (3 movs.) (1773) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil.

No. 27, G major, K. 199 (3 movs.) (1773) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4.

No. 28, C major, K. 200 (4 movs.) (1774) – i: Nil, ii: butterfly, 1:1, min. chord progression, bars 1–4, iii: Nil, iv: Nil.

No. 29, A major, K. 201 (4 movs.) (1774) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: butterfly (Meyer), 1:1, max. chord progression, bars 1–4.

No. 30, D major, K. 202 (4 movs.) (1774) – i: butterfly, 1:1, min. chord progression, bars 1–4, ii: Nil, iii: Nil, iv: Nil.

No. 31, D major, K. 297 (3 movs.) (1778) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 2–5.

No. 32, G major, K. 318 (3 movs.) (1779) – i: Nil, ii: Nil, iii: Nil.

No. 33, B♭ major, K. 319 (4 movs.) (1779) – i: butterfly, 3:1, max. chord progression, bars 2–9, ii: Nil, iii (Menuetto): butterfly, 1:1, max. chord progression, bars 5–8, iv: butterfly, 3:1, max. chord progression, bars 280–287.

No. 34, C major, K. 338 (3 movs.) (1780) – i: Nil, ii: butterfly, 3:1, max. chord progression, bars 1–8, iii: Nil.

No. 35, D major, K. 385 (4 movs.) (1782) – i: Nil, ii: Nil, iii: butterfly, 3:1, min. chord progression, bars 1–8, iv: Nil.

No. 36, C major, K. 425 (4 movs.) (1783) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–8.

No. 37, G major, K. 444 (3 movs.) (1784) – i: Nil, ii: Nil, iii: Nil.

No. 38, D major, K. 504 (3 movs.) (1786) – i: Nil, ii: Nil, iii: Nil.

No. 39, E♭ major, K.543 (4 movs.) (1788) – i: butterfly (Meyer), 3:1, max. chord progression, bars 26–33, ii: Nil, iii: (Menuetto), butterfly, 1:1, max. chord progression, bars 1–8. iv: Nil.

No. 40, G minor, K. 550 (4 movs.) (1788) – i: butterfly, 3:1, min. chord progression, bars 2–9, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

No. 41, C major, K. 551 (4 movs.) (1788) – i: butterfly, 3:1, max. chord progression, bars 1–8, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: butterfly, 3:1, min. chord progression, bars 1–8, iv: butterfly, 1:1, max. chord progression, bars 9–12.

A.4 Mozart String Quartets

No. 1 in G major, K. 80 (4 movs.) (1772) – i: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, ii: Nil, iii: Nil, iv: Nil.

No. 2 in D major, K. 155 (3 movs.) (1772) – i: Nil, ii: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, iii: Nil.

No. 3 in G major, K. 156 (3 movs.) (1772) – i: butterfly (Meyer), 1:1, max. chord progression, bars 1–8, ii: Nil, iii: Nil.

No. 4 in C major, K. 157 (3 movs.) (1772) – i: butterfly, 1:1, max. chord progression, bars 1–4, ii: Nil, iii: Nil.

No. 5 in F major, K. 158 (3 movs.) (1772) – i: butterfly (Aprile), 1:1, max. chord progression, bars 13–20, ii: Nil, iii: Nil.

No. 6 in B♭ major, K. 159 (3 movs.) (1774) – i: Nil, ii: Nil, iii: Nil.

No. 7 in E♭ major, K. 160 (3 movs.) (1773) – i: Nil, ii: Nil, iii: Nil.

No. 8 in F major, K. 168 (4 movs.) (1773) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 9 in A major, K. 169 (4 movs.) (1773) – i: butterfly, 1:1, min. chord progression, bars 12–15, ii: Nil, iii: Nil, iv: Nil.

No. 10 in C major, K. 170 (4 movs.) (1773) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

No. 11 in E♭ major, K. 171 (4 movs.) (1773) – i: Nil, ii: Nil, iii: butterfly, 1:1, min. chord progression, bars 1–4, iv: Nil.

No. 12 in B♭ major, K. 172 (4 movs.) (1773) – i: Nil, ii: butterfly (Aprile), 1:1, max. chord progression, bars 1–2, iii (Trio): butterfly, 1:1, max. chord progression, bars 31–34, iv: Nil.

No. 13 in D minor, K. 173 (4 movs.) (1773) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 14 in G major, K. 387 (4 movs.) (1782) – i: butterfly, 3:1, min. chord progression, ii: butterfly, 1:1, max. chord progression, bars 21–24, iii: Nil, iv: Nil.

No. 15 in D minor, K. 421 (4 movs.) (1783) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 16 in E♭ major, K. 428 (4 movs.) (1783) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 17 in B♭ major, K. 458 (4 movs.) (1784) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 18 in A major, K. 464 (4 movs.) (1785) – i: butterfly, 3:1, max. chord progression, bars 1–8, ii: butterfly, 1:1 max. chord progression, bars 1–4, iii: Nil, iv: Nil.

No. 19 in C major, K. 465 (4 movs.) (1785) – i: butterfly, 1:1, min. chord progression, bars 23–26, ii: Nil, iii (Trio): butterfly (Meyer), 3:1, max. chord progression, bars 1–8, iv: Nil.

No. 20 in D major, K. 499 (4 movs.) (1786) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 21 in D major, K. 575 (4 movs.) (1789) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 22 in B♭ major, K. 589 (4 movs.) (1790) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 23 in F major, K. 590 (4 movs.) (1790) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A.5 Beethoven Piano Sonatas

E♭ major, WoO 47, No. 1 (3 movs.) (1783) – i: butterfly, 1:1, max. chord progression, bars 4–5, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil.

F minor, WoO 47, No. 2 (3 movs.) (1782) – i: Nil, ii: Nil, iii: Nil.

D major, WoO 47, No. 3 (3 movs.) (1783) – i: butterfly, 1:1, max. chord progression, bars 8–10, ii: Nil, iii: Nil.

F minor, Op. 2, No. 1 (4 movs.) (1793–95) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A major, Op. 2, No. 2 (4 movs.) (1795) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, min. chord progression, bars 13–16.

C major, Op. 2, No. 3 (4 movs.) (1795) – i: butterfly, 1:1, max. chord progression, bars 1–4, ii: Nil, iii: Nil, iv: Nil.

G minor, Op. 49, No. 1 (2 movs.) (1797) – i: Nil, ii: Nil.

G major, Op. 49, No. 2 (2 movs.) (1796) – I: butterfly, 1:1, max. chord progression, bars 21–24, II: Nil.

E♭ major, Op. 7, (4 movs.) (1796–97) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

C minor, Op. 10, No. 1 (3 movs.) (1795–98) – i: butterfly (Meyer), 1:1, max. chord progression, bars 56–63, ii: Nil, iii: Nil.

F major, Op. 10, No. 2 (3 movs.) (1797–98) – i: butterfly, 3:1, max. chord progression, bars 19–26, ii: Nil, iii: Nil.

D major, Op. 10, No. 3 (4 movs.) (1797–98) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 55–62, iv: Nil.

C minor, Op. 13 (3 movs.) (1796–99) – i: butterfly, 1:1, max. chord progression, bars 53–60, ii: Nil, iii: Nil.

E major, Op. 14, No. 1 (3 movs.) (1798) – i: Nil, ii: Nil, iii: Nil.

G major, Op. 14, No. 2 (3 movs.) (1798) – i: Nil, ii: Nil, iii: Nil.

B♭ major, Op. 22 (4 movs.) (1800) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–4.

A♭ major, Op. 26 (3 movs.) (1800–01) – i: Nil, ii: Nil, iii: Nil.

E♭ major, Op. 27, No. 1 (3 movs.) (1801) – i: butterfly (Meyer), 1:1, max. chord progression, bars 37–40, ii: Nil, iii: Nil.

C♯ minor, Op. 27, No. 2 (3 movs.) (1801) – i: Nil, ii: Nil, iii: Nil.

D major, Op. 28 (4 movs.) (1801) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 4–8, iv: Nil.

G major, Op. 31, No. 1 (3 movs.) (1803) – i: Nil, ii: Nil, iii: Nil.

D minor, Op. 31, No. 2 (3 movs.) (1802) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–8.

E♭ major, Op. 31, No. 3 (4 movs.) (1802) – i: Nil, ii: Nil, iii (Menuetto): butterfly (Meyer), 1:1, max. chord progression, bars 1–4, iv: Nil.

C major, Op. 53 (2 movs.) (1804) – i: Nil, ii: Nil.

F major, Op. 54 (2 movs.) (1804) – i: Nil, ii: Nil.

F minor, Op. 57 (3 movs.) (1804–06) – i: Nil, ii: Nil, iii: Nil.

F♯ minor, Op. 78 (2 movs.) (1809) – i: Nil, ii: Nil.

G major, Op. 79 (3 movs.) (1809) – i: Nil, ii: Nil, iii: Nil.

E♭ major, Op. 81a (3 movs.) (1809) – i: Nil, ii: Nil, iii: Nil.

E minor, Op. 90 (2 movs.) (1814) – i: Nil, ii: butterfly, 1:1, min. chord progression, bars 1–4.

A major, Op. 101 (3 movs.) (1815–16) – i: Nil, ii: Nil, iii: Nil.

B♭ major, Op. 106 (4 movs.) (1817–18) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

E major, Op. 109 (3 movs.) (1820) – i: Nil, ii: Nil, iii: Nil.

A♭ major, Op. 110 (4 movs.) (1821) – i: butterfly, 1:1, max. chord progression, bars 5–8, ii: Nil, iii: Nil, iv: Nil.

C minor, Op. 111 (2 movs.) (1822) – i: Nil, ii: Nil.

A.6 Beethoven String Quartets

No. 1 in F major, Op. 18, No. 1 (4 movs.) (1799) – i: Nil, ii: Nil, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–4.

No. 2 in G major, Op. 18, No. 2 (4 movs.) (1799) – i: Nil, ii: Nil, iii (Scherzo): 3:1, min. chord progression, bars 1–8, iv: butterfly, 3:1, min. chord progression, bars 140–147.

No. 3 in D major, Op. 18, No. 3 (4 movs.) (1798/99) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 4 in C minor, Op. 18, No. 4 (4 movs.) (1799–1800) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 5 in A major, Op. 18, No. 5 (4 movs.) (1799) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

No. 6 in B♭ major, Op. 18, No. 6 (4 movs.) (1800) – i: Nil, ii: butterfly, 1:1, max. chord progression, bars 1–4, iii: Nil, iv: Nil.

No. 7 in F major, Op. 59, No. 1 (4 movs.) (1806) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 8 in E minor, Op. 59, No. 2 (4 movs.) (1806) – i: Nil, ii: Nil, iii: butterfly, 1:1, max. chord progression, bars 1–4, iv: Nil.

No. 9 in C major, Op. 59, No. 3 (4 movs.) (1806) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 10 in E♭ major, Op. 74 (4 movs.) (1809) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 11 in F minor, Op. 95 (5 movs.) (1810) – i: Nil, ii: Nil, iii: Nil, iv: Nil, v: Nil.

No. 12 in E♭ major, Op. 127 (4 movs.) (1825) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 13 in B♭ major, Op. 130 (6 movs.) (1825) – i: Nil, ii: Nil, iii: Nil, iv: Nil, v: Nil, vi: Nil.

No. 14 in C♯ minor, Op. 131 (7 movs.) (1826) – i: Nil, ii: Nil, iii: Nil, iv: butterfly (Meyer), 1:1, max. chord progression, bars 1–4, v: Nil, vi: Nil, vii: Nil.

No. 15 in A minor, Op. 132 (5 movs.) (1825) – i: Nil, ii: Nil, iii: Nil, iv: Nil, v: Nil.

No. 16 in F major, Op. 135 (4 movs.) (1826) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A.7 Schubert String Quartets

No. 1, D. 18 (4 movs.) (1812) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 2, D. 32 (2 movs.) (1812) – i: Nil, ii: Nil.

No. 3, D. 36 (4 movs.) (1812) – i: Nil, ii: Nil, iii: butterfly, 3:1, max. chord progression, bars 1–8, iv: Nil.

No. 4, D. 46 (4 movs) (1813) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 5, D. 68 (2 movs.) (1813) – i: Nil, ii: Nil.

No. 6, D. 74 (4 movs.) (1813) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 7, D. 94 (4 movs.) (1814) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 8 in B♭ major, D. 112 (4 movs.) (1814) – i: Nil, ii: Nil, iii (Menuetto): butterfly, 1:1, max. chord progression, bars 1–8, iv: Nil.

No. 9, D. 173 (4 movs.) (1815) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 10, D. 87 (4 movs. (1813) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 11 in E major, D. 353 (4 movs.) (1816) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 12 in C minor, D. 703 (1 mov.) (1820) – i: Nil.

No. 13 in A minor, D. 804 (4 movs.) (1824) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 14 in D minor, D. 810 (4 movs.) (1824) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 15 in G major, D. 887 (4 movs.) (1826) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

A.8 Schubert Piano Sonatas

No. 1 in E major, D. 157 (3 movs.) (1815) – i: butterfly, 3:1, max. chord progression, bars 46–50, ii: butterfly, 1:1, max. chord progression, bars 1–2, iii: Nil.

No. 2 in C major, D. 279 (3 movs.) (1815) – i: Nil, ii: Nil, iii: Nil.

No. 3 in E major, D. 459 (2 movs.) (1816) – i: butterfly, 1:1, max. chord progression, bars 1–4, ii: Nil.

No. 4 in A minor, D. 537 (3 movs.) (1817) – i: Nil, ii: Nil, iii: Nil.

No. 5 in A♭ major, D. 557 (3 movs.) (1817) – i: Nil, ii: Nil, iii: Nil.

No. 6 in E minor, D. 566 (3 movs.) (1817) – i: Nil, ii: Nil, iii: Nil.

No. 7 in E♭ major, D. 568 (4 movs.) (1817) – i: butterfly, 1:1, max. chord progression, bars 41–44, ii: Nil, iii: Nil, iv: butterfly, 1:1, max. chord progression, bars 1–2.

No. 8 in F♯ minor, D. 571 (1 mov.) (1817) – i: Nil.

No. 9 in B major, D. 575 (4 movs.) (1817) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 10 in C major, D. 613 (2 movs.) (1818) – i: Nil, ii: Nil.

No. 11 in F minor, D. 625 (4 movs.) (1818) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 12 in C♯ minor, D. 655 (1 mov.) (1819) – i: Nil.

No. 13 in A major, D. 664 (3 movs.) (1819) – i: Nil, ii: Nil, iii: Nil.

No. 14 in A minor, D. 784 (3 movs.) (1823) – i: Nil, ii: Nil, iii: Nil.

No. 15 in C major, D. 840 (4 movs.) (1825) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 16 in A minor, D. 845 (4 movs.) (1825) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 17 in D major, D. 850 (4 movs.) (1826) – i: Nil, ii: Nil, iii (Scherzo): butterfly, 1:1, max. chord progression, bars 50–57, iv: Nil.

No. 18 in G major, D. 894 (4 movs.) (1826) – i: Nil, ii: Nil, iii (Trio): butterfly, 1:1, max. chord progression, bars 3–6, iv: Nil.

No. 19 in C minor, D. 958 (4 movs.) (1828) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 20 in A major, D. 959 (4 movs.) (1828) – i: Nil, ii: Nil, iii: Nil, iv: Nil.

No. 21 in B♭ major, D. 960 (4 movs.) (1828) – i: Nil, ii: butterfly, 1:1, min. chord progression, bars 1–4, iii: Nil, iv: Nil.

A.9 Chopin Mazurkas

F♯ minor, Op. 6, No. 1 (1830) – Nil.

C♯ minor, Op. 6, No. 2 (1830) – Nil.

E major, Op. 6, No. 3 (1830) – Nil.

E♭ minor, Op. 6, No. 4 (1830) – Nil.

B♭ major, Op. 7, No. 1 (1824–31) – Nil.

A minor, Op. 7, No. 2 (1824–31) – Nil.

F minor, Op. 7, No. 3 (1824–31) – Nil.

A♭ major, Op. 7, No. 4 (1824–31) – Nil.

C major, Op. 7, No. 5 (1824–31) – Nil.

B♭ major, Op. 17, No. 1 (1831–33) – Nil.

E minor, Op. 17, No. 2 (1831–33) – Nil.

A♭ major, Op. 17, No. 3 (1831–33) – Nil.

A minor, Op. 17, No. 4 (1831–33) – Nil.

G minor, Op. 24, No. 1 (1834–35) – Nil.

C major, Op. 24, No. 2 (1834–35) – Nil.

A♭ major, Op. 24, No. 3 (1834–35) – Nil.

B♭ minor, Op. 24, No. 4 (1834–35) – Nil.

C minor, Op. 30, No. 1 (1836–37) – Nil.

B minor, Op. 30, No. 2 (1836–37) – Nil.

D♭ major, Op. 30, No. 3 (1836–37) – Nil.

C♯ minor, Op. 30, No. 4 (1836–37) – Nil.

G♯ minor, Op. 33, No. 1 (1837–38) – Nil.

D major, Op. 33, No. 2 (1837–38) – butterfly, 1:1, max. chord progression, bars 5–8.

C major, Op. 33, No. 3 (1837–38) – Nil.

B minor, Op. 33, No. 4 (1837–38) – Nil.

C♯ minor, Op. 41, No. 1 (1838–39) – Nil.

E minor, Op. 41, No. 2 (1838–39) – Nil.

B major, Op. 41, No. 3 (1838–39) – Nil.

A♭ major, Op. 41, No. 4 (1838–39) – Nil.

G major, Op. 50, No. 1 (1841–42) – Nil.

A♭ major, Op. 50, No. 2 (1841–42) – Nil.

C♯ minor, Op. 50, No. 3 (1841–42) – Nil.

B major, Op. 56, No. 1 (1843) – Nil.

C major, Op. 56, No. 2 (1843) – Nil.

C minor, Op. 56, No. 3 (1843) – Nil.

A minor, Op. 59, No. 1 (1849) – butterfly, 1:1, min. chord progression, bars 1–4.

A♭ major, Op. 59, No. 2 (1849) – Nil.

F♯ minor, Op. 59, No. 3 (1849) – Nil.

B major, Op. 63, No. 1 (1846) – Nil.

F minor, Op. 63, No. 2 (1846) – Nil.

C♯ minor, Op. 63, No. 3 (1846) – Nil.

G major, Op. 67, No. 1 (1830–1848) – Nil.

G minor, Op. 67, No. 2 (1830–1848) – Nil.

C major, Op. 67, No. 3 (1830–1848) – Nil.

A minor, Op. 67, No. 4 (1830–1848) – butterfly, 1:1, max. chord progression, bars 34–37.

C major, Op. 68, No. 1 (1827–49) – Nil.

A minor, Op. 68, No. 2 (1827–49) – Nil.

F major, Op. 68, No. 3 (1827–49) – Nil.

F minor, Op. 68, No. 3 (1827–49) – Nil.

B♭ major, 'Wolowska' (1832) – Nil.

G major, B. 16 (1826) – Nil.

B♭ major, B. 16 (1826) – Nil.

D major, B. 31 (1829) – Nil.

C major, B. 82 (1833) – Nil.

A minor, B. 134 (?) – Nil.

A minor, B. 140 (?) – Nil.

A.10 Chopin Nocturnes

B♭ minor, Op. 9, No. 1 (1830–31) – Nil.

E♭ major, Op. 9, No. 2 (1830–31) – Nil.

B major, Op. 9, No. 3 (1830–31) – Nil.

F major, Op. 15, No. 1 (1830–33) – Nil.

F♯ major, Op. 15, No. 2 (1830–33) – Nil.

G minor, Op. 15, No. 3 (1830–33) – Nil.

C♯ minor, Op. 27, No. 1 (1836) – Nil.

D♭ major, Op. 27, No. 2 (1836) – Nil.

B major, Op. 32, No. 1 (1836–37) – Nil.

A♭ major, Op. 32, No. 2 (1836–37) – Nil.

G minor, Op. 37, No. 1 (1840) – Nil.

G major, Op. 37, No. 2 (1840) – Nil.

C minor, Op. 48, No. 1 (1841) – Nil.

F♯ minor, Op. 48, No. 2 (1841) – Nil.

F minor, Op. 55, No. 1 (1842–44) – Nil.

E♭ major, Op. 55, No. 2 (1842–44) – Nil.

B major, Op. 62, No. 1 (1846) – Nil.

E major, Op. 62, No. 2 (1846) – Nil.

E minor, Op. 72, No. 1 (1827) – Nil.

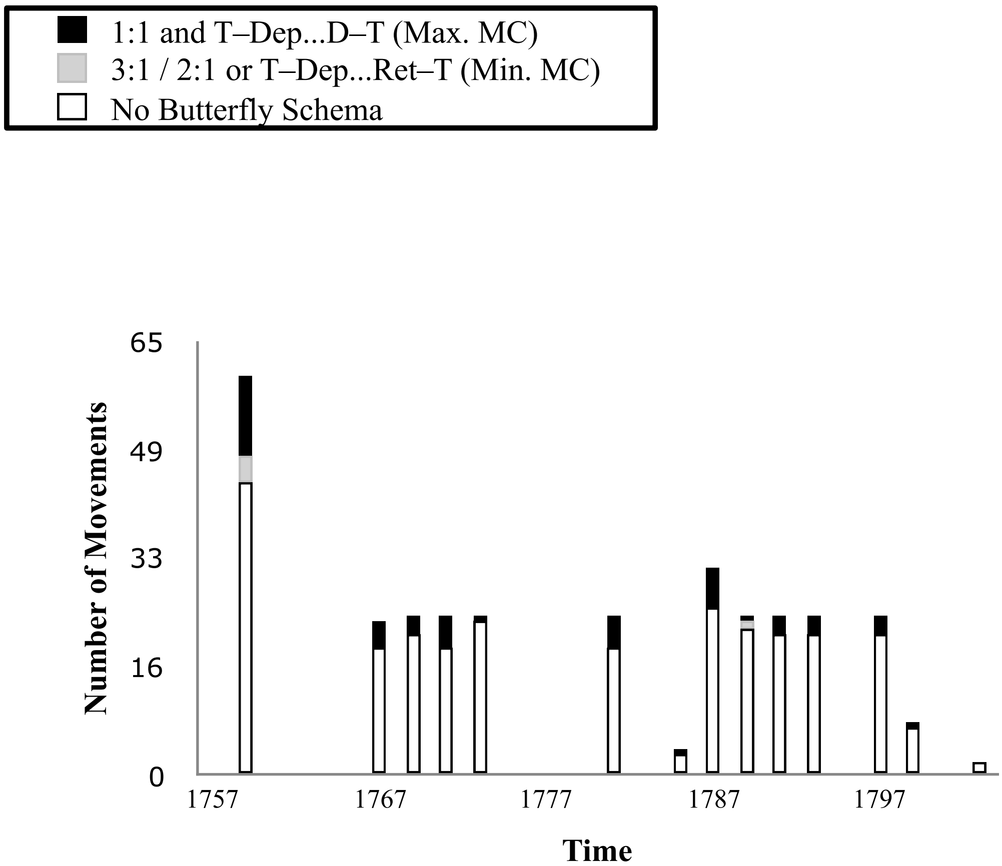
C♯ minor, B. 49 (1830) – Nil.

C minor, B. 108 (1837) – Nil.

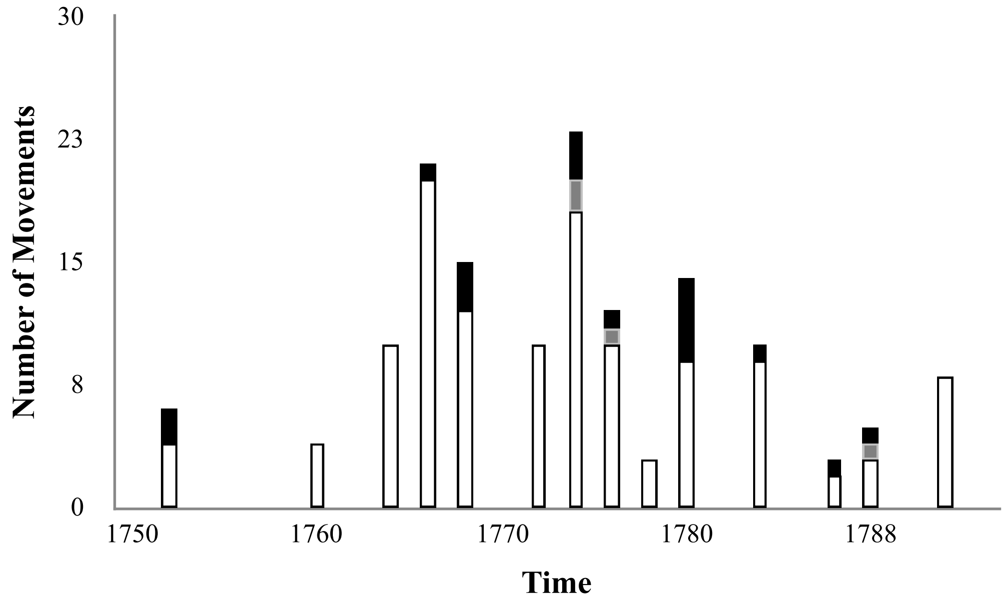
Nocturne oubliée (1833) – Nil.

Appendix B

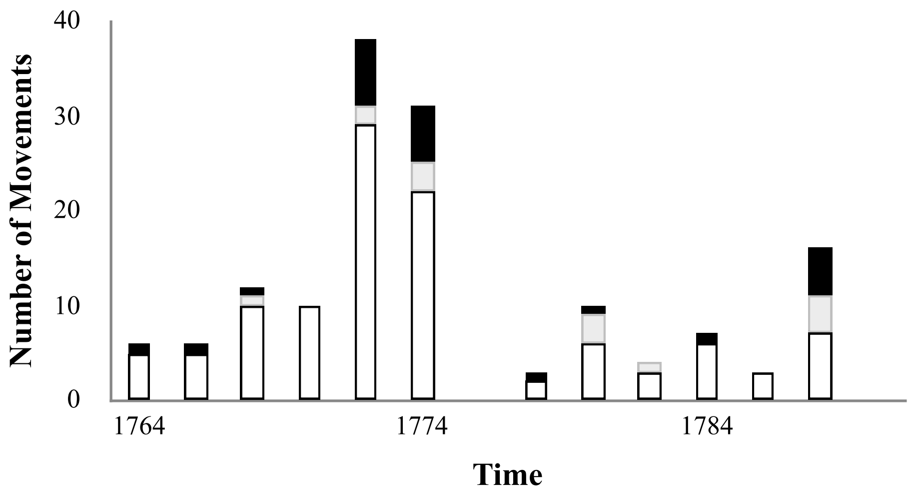
Charts Showing the Type and Distribution of Butterfly Schemata in Corpora of Classical and Romantic Composers



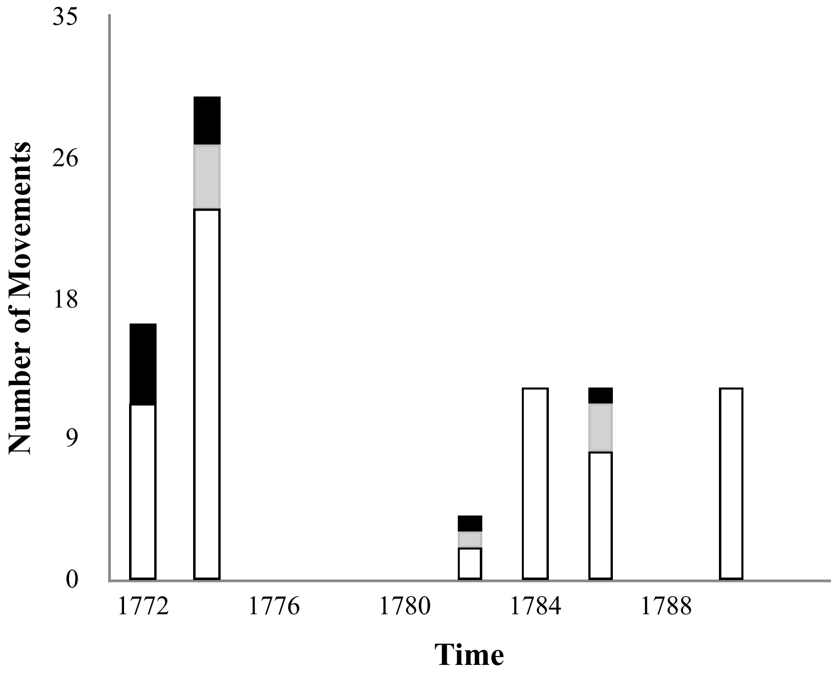
B.1 Butterfly schemata in Haydn string quartets.



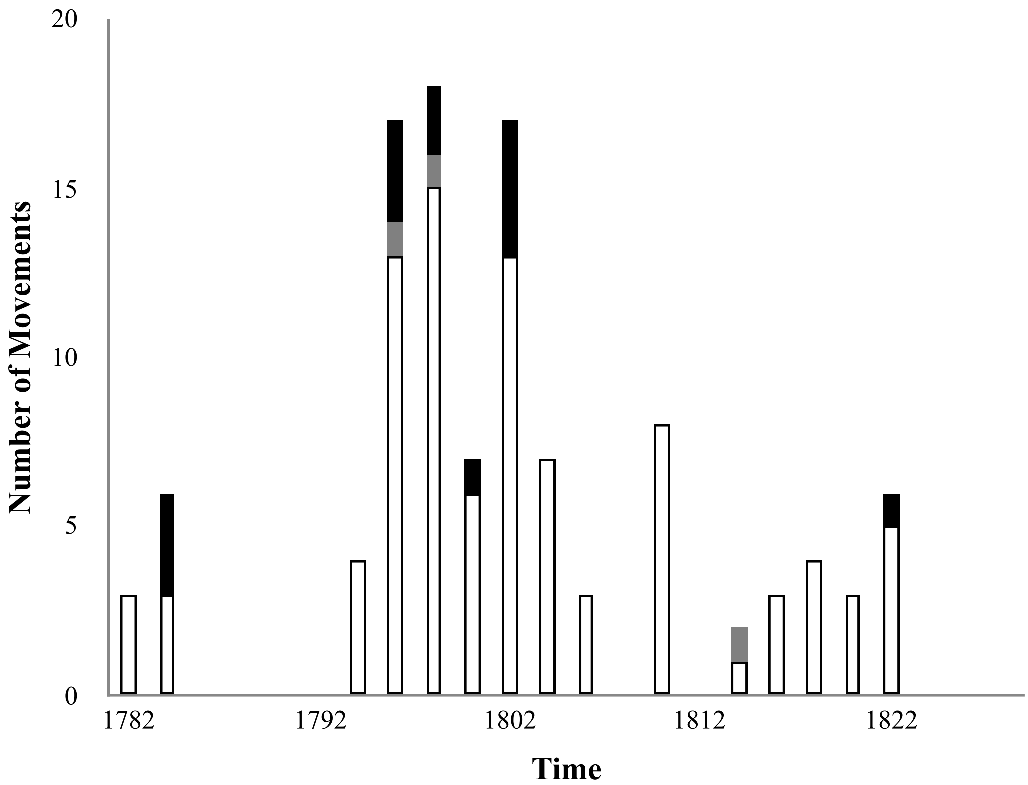
B.2 Butterfly schemata in Haydn piano sonatas.



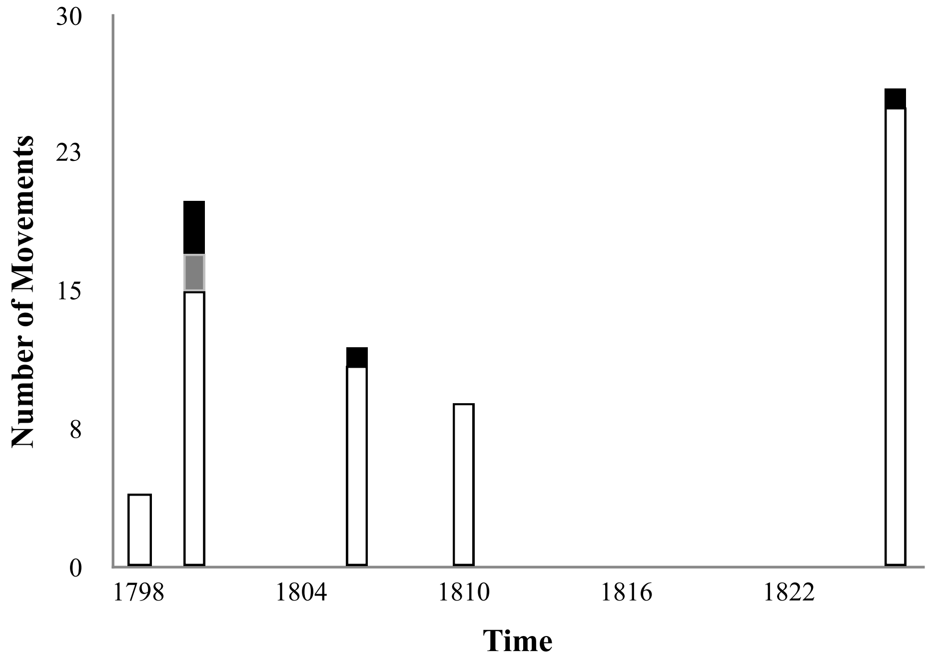
B.3 Butterfly schemata in Mozart symphonies.



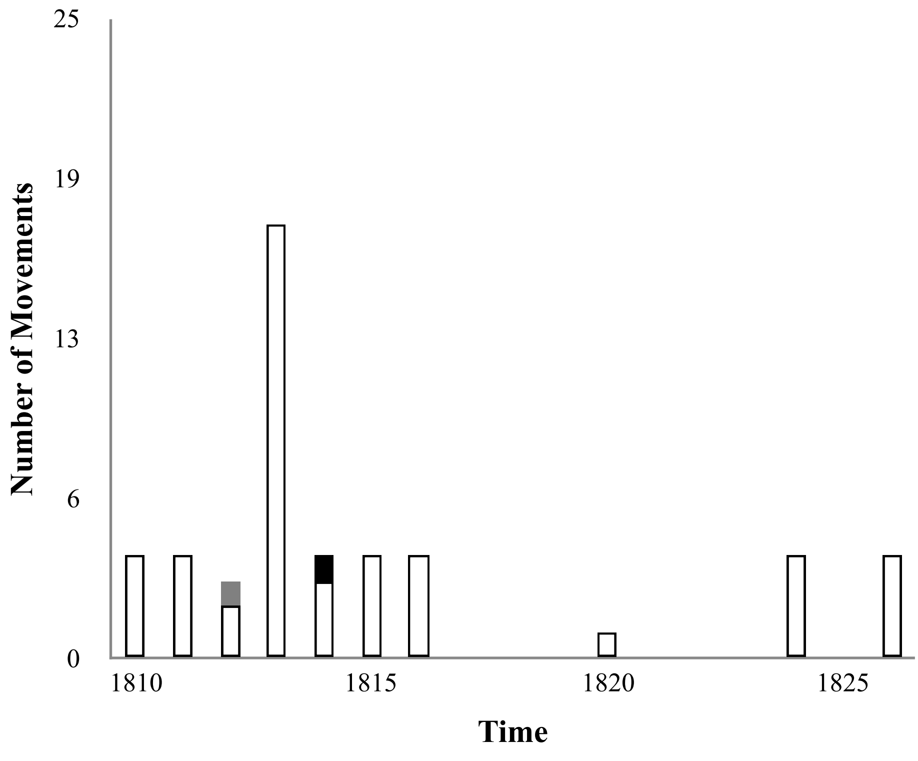
B.4 Butterfly schemata in Mozart string quartets.



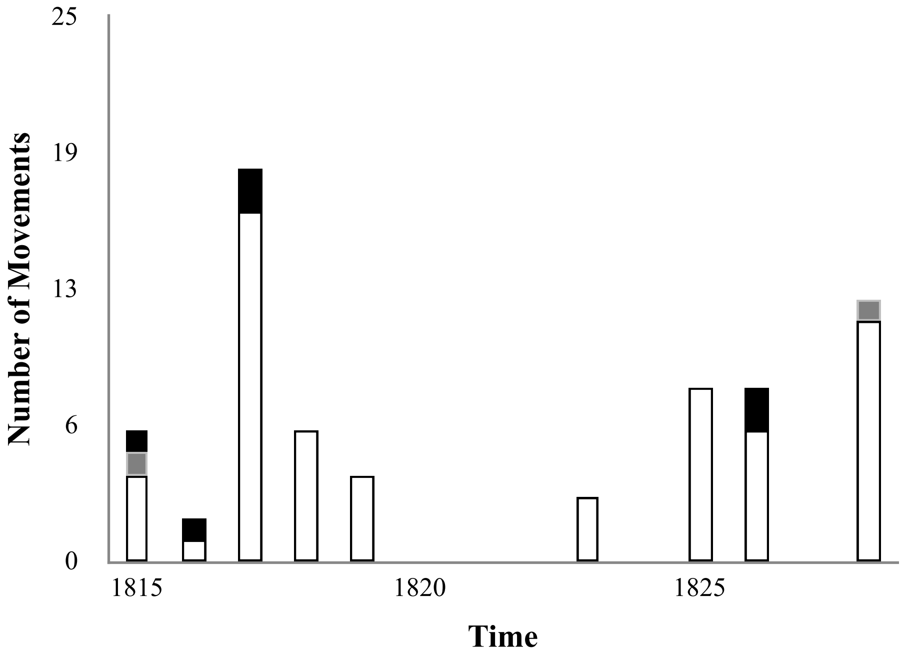
B.5 Butterfly schemata in Beethoven piano sonatas.



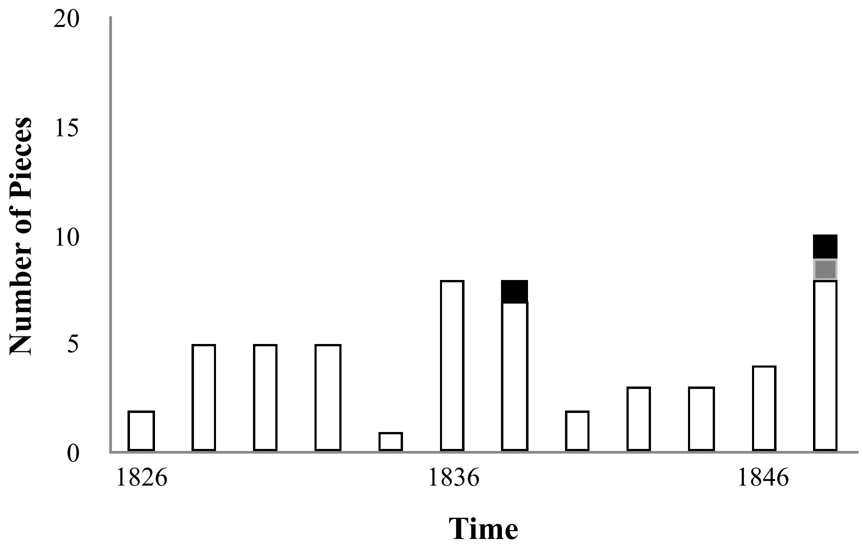
B.6 Butterfly schemata in Beethoven string quartets.



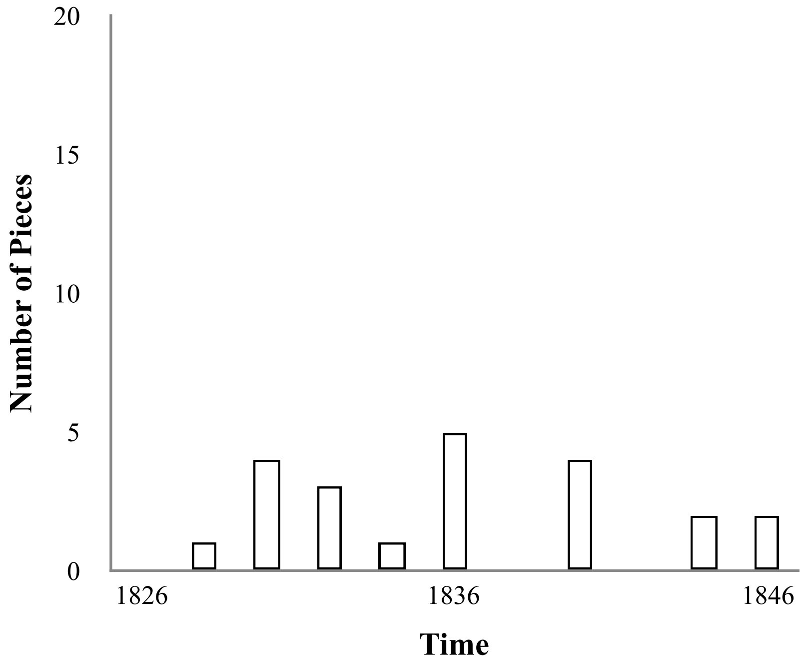
B.7 Butterfly schemata in Schubert string quartets.



B.8 Butterfly schemata in Schubert piano sonatas.



B.9 Butterfly schemata in Chopin mazurkas.



B.10 Butterfly schemata in Chopin nocturnes.

Appendix C

An Independent-Samples T-Test Comparing the Means of Butterfly Schemata Found in the Classical-Period Samples and Romantic-period Samples

An independent-samples t-test was performed to test whether the difference between the means of the Classical- and Romantic-period samples was statistically significant. Butterfly schemata were more frequently found in the Classical-period sample (M = 21.8, SD = 4.0) than in the Romantic period (M = 6.6, SD = 4.6), t = 6.09, p = 0.0001. The test showed that the difference between the means of the samples was unlikely to have occurred by chance (p = 0.0001).

C.1 Percentage of butterfly schemata found in Classical- and Romantic-period samples.

Classical-period samples	**Romantic-period samples**
Hadyn Quartets (24.1 %)	Post-1800 Beethoven Sonatas (11.7%)
Haydn Sonatas (15.3 %)	Post-1800 Beethoven Quartets (6.8 %)
Mozart Symphonies (26 %)	Schubert Quartets (3.8 %)
Mozart Quartets (19.8 %)	Schubert Sonatas (11.9 %)
Pre-1800 Beethoven Sonatas (20.8 %)	Chopin Mazurkas (5.4 %)
Pre-1800 Beethoven Quartets (25 %)	Chopin Nocturnes (0 %)

C.2 Comparison of means, SD, and variance in the Classical- and Romantic-period samples.

Comparison of means				
	Sample Size	Mean	Standard Deviation	Variance
Classical sample	6	21.83333	4.01032	16.08267
Romantic sample	6	6.6	4.62558	21.396

C.3 An Independent-samples t-test comparing Classical- and Romantic-period samples.

T-test assuming equal variances (homoscedastic)	
Degrees of Freedom	10
Hypothesised Mean Difference	0
Pooled Variance	18.73933
Test Statistics	6.09507

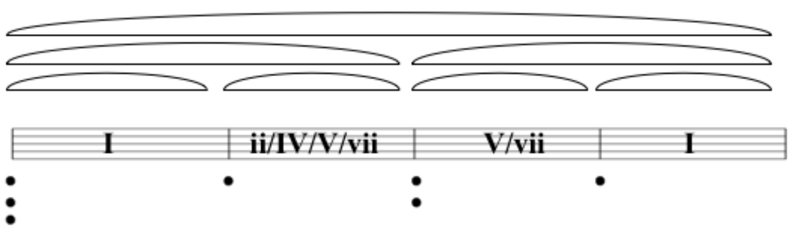
C.4 Probability values based on the test statistics.

Two-tailed distribution	
p-level	0.00012 (Critical Value 5%)
One-tailed distribution	
p-level	0.00006 (Critical Value 5%)

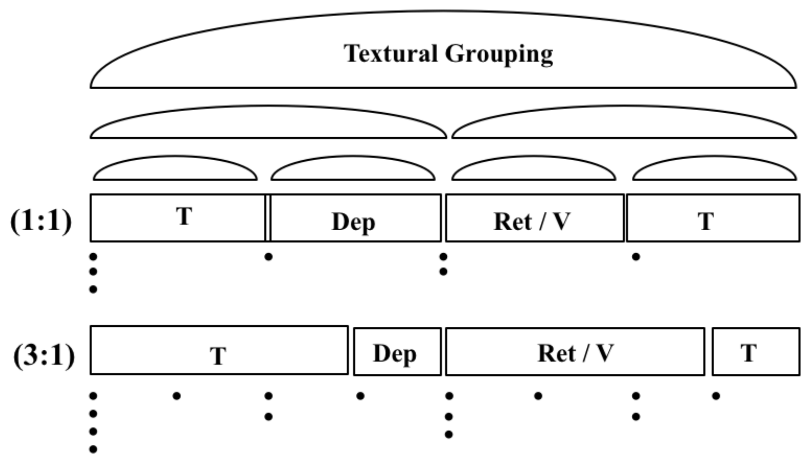
Appendix D

Statement of Models and Rules Used to Define and Validate the Butterfly Schema

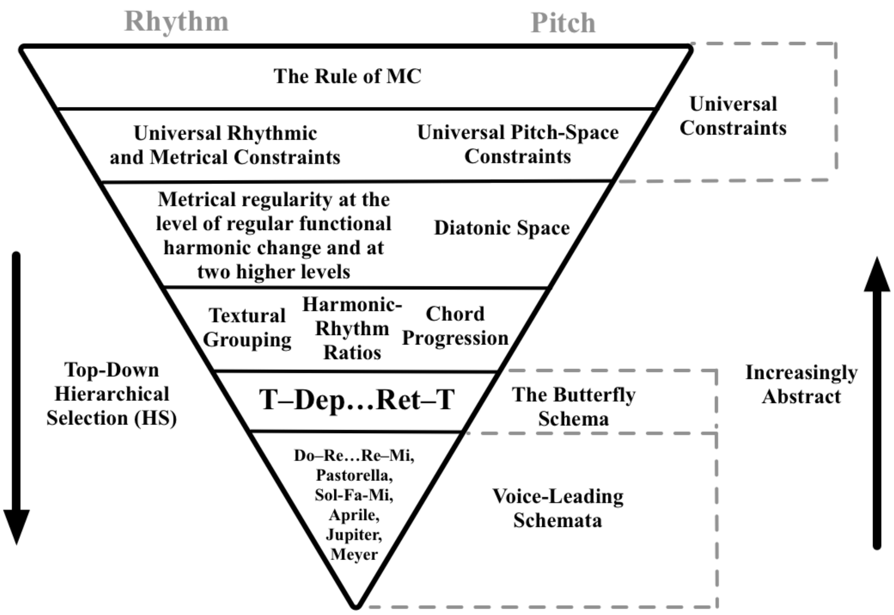
1. Preliminary model of the butterfly schema.



2. Final multiparametrically congruent model of the butterfly schema.



3. Top-down hierarchical selection (HS) model of the butterfly schema.



4. List of rules used to define and validate butterfly schemata.

(i) Functional chords are in a single key and key area.

(ii) The T regions include tonic chords, $\delta = 0$.

(iii) The minimal hierarchical distance between functional chords in the Dep and Ret regions and the Ts is $\delta = 5$.

(iv) The maximal hierarchical distance between functional chords in the Dep and Ret regions and the Ts is $\delta = 8$.

(v) T–Dep…Ret–T_{min}.

(vi) T–Dep…V–T_{max}.

(vii) Chords that are not the functional chords of Dep and Ret regions (including non-diatonic, passing, interpolated and appoggiatura chords) are valid even if they are significant

in the time-span reduction of those regions, occurring on the strongest beats of the regions, provided their resolution is on the next strongest beat in the same region as the functional chord.

(viii) Non-functional chords are valid in time-span significant positions in T regions provided that their resolution is to a tonic chord in those regions (on a weak or strong beat).

(ix) Tonic chords are not valid on weak or strong beats in the Dep and Ret regions.

(x) Regular, duple and hierarchical textural grouping occurs at the level of regular functional harmonic change and at two immediately higher levels of the metrical structure.

(xi) The harmonic-rhythm ratios 1:1 and 3:1 are uniparametrically congruent features.

(xii) A 1:1 harmonic-rhythm ratio is maximally uniparametrically congruent.

(xiii) A 3:1 harmonic-rhythm ratio is minimally uniparametrically congruent.

(*xiv*) *A minimally congruent butterfly schema =*
(*rule x*) + ((*rule v + rule xiii*) ***or*** (*rule v + rule xii)* ***or*** (*rule vi + rule xiii*)).

(*xv*) *A maximally congruent butterfly schema = rule x + rule vi + rule xii.*

(xvi) Features of the butterfly schema and the Classical grammar are culturally selected against each other in a top-down hierarchical order, with the rule of MC governing each level of the hierarchy.

References

Agawu, V.K. (1991). *Playing with signs*: *A semiotic interpretation of classic music*. Princeton: Princeton University Press.

Baldwin, J.M. (1896). A new factor in evolution. *The American Naturalist*, (*30*)354, 441–451.

Baroni, M., Maguire, S., Drabkin, W. (1983). The concept of musical grammar. *Music Analysis*, (*2*)2, 175–208.

Baum, W.A. (1994). *Understanding behaviorism: Behavior, culture, and evolution*. Malden, MA: Blackwell Publishing.

Bauman, A. (1952). The thematic process in music by Rudolph Reti. Review. *Journal of the American Musicological Society,* (*5*)2, 139–141.

Benjamin, W. (1984). A theory of musical meter. *Music Perception*, (*1*), 355–413.

Berlin, B. and Kay, P. (1969). *Basic color terms*: *Their universality and evolution*. Berkeley: University of California Press.

Bernstein, L. (1976). *The unanswered question*: *Six talks at Harvard*. Cambridge, Massachusetts: Harvard University Press.

Berry, D.C. (2005). Schenkerian theory in the United States: A review of its establishment and a survey of current research topics. *Zeitschrift der Gesellschaft der Musiktheorie*, (*2*)2–3, 101–137.

Berry, W. (1976). *Structural functions in music*. Englewood Cliffs, NJ: Prentice-Hall.

Bertalanffy, L.v. (1968). *General system theory*: *Foundations*, *development*, *applications*. New York: George Braziller.

Bharucha, J.J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, (*5*), 1–30.

______. (1991). Pitch, harmony and neural nets: A psychological perspective. In P. M. Todd & D. G. Loy (Eds.), *Music and connectionism* (pp. 84–99). Cambridge, MA: MIT Press.

Blackmore, S. (1999). *The meme machine*. Oxford: Oxford University Press.

Bod, R. (2001b). A memory-based model for music analysis. *Proceedings of the International Computer Music Conference*, *ICMC* 2001, Havana, Cuba.

______. (2001c). Memory-based models of melodic analysis: Challenging the gestalt principles. *Journal of New Music Research*, (*31*)1.

______. (2002). A general parsing model for music and language. *Proceedings of the Second International Conference*, *ICMAI 2002*, *Edinburgh*, *UK*, 5–17.

Bregman, A.S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.

Bruce, V., Green, P. & Georgeson, M. (1996). *Visual perception*: *physiology, psychology and ecology* (3rd ed.). New York: Psychology Press.

Byros, V. (2015). Hauptruhepuncte des geistes: Punctuation schemas and the late-eighteenth-century sonata. In M. Neuwirth, and P. Berge (Eds.), *What is a Cadence*: *Theoretical and Analytical Perspectives on Cadences in the Classical Repertoire* (pp. 215–251). Leuven: Leuven University Press.

______. (2009a). *Foundations of tonality as situated cognition, 1730–1830*: *An enquiry into the culture and cognition of eighteenth-century tonality with Beethoven's Eroica symphony as a case study.* Ph.D. diss., Yale University.

______. (2009b). Towards an "archaeology" of hearing: Schemata and eighteenth-century consciousness. *Musica Humana*, (*1*)2, 235–306.

______. (2012a). Unearthing the past: Theory and archaeology in Robert Gjerdingen's *music in the galant style*. *Music Analysis*, (*31*)1, 112–123.

______. (2012b). Meyer's anvil: Revisiting the schema concept. *Music Analysis*, (*31*)3, 273–346.

Cambouropoulus, E. (1998). *Towards a general computational theory of musical structure*. Ph.D. diss., University of Edinburgh, UK.

Carlyon, R.P., Cusack, R., Foxton, J.M., and Robertson, I.H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology*: *Human Perception and Performance*, (*27*), 115–127.

Caplin, W. E. (1998). *Classical form*: *A theory of formal functions for the music of Haydn, Mozart, and Beethoven*. Oxford: Oxford University Press.

Cavett-Dunsby, E. (1990). A classic turn of phrase: Music and the psychology of convention by Robert Gjerdingen. Review. (9)1, 79–85.

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies* (*2*)3, 200–219.

Chatter, N. (1999). The search for simplicity: A fundamental cognitive principle?. *The Quarterly Journal of Experimental Psychology*, (*52*)2, 273–302.

Chatter and Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science?. *Trends in Cognitive Sciences*, (*7*)1, 19–22.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

______. (1959). Verbal behavior by B.F. Skinner. Review. *Language*, (*35*)1, 26–58.

______. (1965). *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.

______. (1966). *Cartesian linguistics*: *A chapter in the history of rationalist thought*. New York: Harper & Row.

______. (1975). Acquisition and use of language. In S. P. Stich (Ed.), *Innate Ideas* (pp. 89–103). Los Angeles: University of California Press.

______. (2000). *The architecture of language*. Oxford: Oxford University Press.

______. (2006). *Language and mind*. Cambridge: Cambridge University Press.

Christensen, T. (2002). Introduction. In T. Chistensen (Ed.), *The Cambridge history of Western music theory* (pp. 1–23). Cambridge: Cambridge University Press.

Clarke, E. (1989). Mind the gap: Formal structures and psychological processes in music. *Contemporary Music Review*, (*3*), 1–13.

Cohn, R. (1996). Maximally smooth cycles, hexatonic systems, and the analysis of late-romantic triadic progressions. *Music Analysis*, (*15*), 9–40.

______. (1997). Neo-Riemannian operations, parsimonious trichords, and their *tonnetz* representations. *Journal of Music Theory*, (*41*), 1–66.

______. (1998). Square dances with cubes. *Journal of Music Theory*, (*42*), 283–296.

______. (1999). As wonderful as star clusters: Instruments for gazing at tonality in Schubert. *Nineteenth Century Music,* (*22*), 213–232.

Cohn, R., and Dempster, D. (1992). Hierarchical unity, plural unities: Toward a reconciliation. In K. Bergeron and P.V. Bohlmann (Eds.), *Disciplining music*: *Musicology and its canons* (pp. 156–181). Chicago: University of Chicago Press.

Cone, E.T. (1968). *Musical form and musical performance*. New York: Norton.

______. (1967). Beyond analysis. *Perspectives of New Music*, (*6*)1, 33–51.

Conklin, D. (2002). Representation and discovery of vertical patterns in music. In C. Anagnostopoulou, M. Ferrand, and A. Smaill (Eds.), *Music and artificial intelligence* (pp. 32–42). Springer-Verlag.

Conklin, D. and Cleary, J.G. (1988). Modeling and generating music using multiple viewpoints. *Proceedings of the First Workshop on AI and Music*, 125–137.

Conklin, D. and Witten, I.H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, (*24*)1, 51–73.

Cope, D. (1991a). *Computers and musical style*. Oxford: Oxford University Press.

______. (1991b). *Recombinant music*: *Using the computer to explore musical style*. Oxford: Oxford University Press.

______. (1996). *Experiments in musical intelligence*. Madison, WI: A-R Editions.

______. (2000b). *The algorithmic composer*. Madison, WI: A-R Editions.

______. (2001). *Virtual music: Computer synthesis of musical style*. Cambridge, Massachusetts: The MIT Press.

Cook, N. (1987a). The perception of large-scale tonal closure. *Music Perception*, (*5*), 197–205.

______. (1987b). *A guide to music analysis*. Oxford: Oxford University Press.

______. (1998). *Music*: *A very short introduction*. Oxford: Oxford University Press.

Cooper, G., and Meyer, L.B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.

Darwin, C. (1859). On the origins of species by means of natural selection. London: John Murray.

Dawkins, R. (1976). *The selfish gene*. New York City: Oxford University Press.

______. (1986). *The blind watchmaker*. New York: W. W. Norton & Company.

______. (1999). Foreword. In S. Blackmore, *The meme machine* (pp. vii–xvii). Oxford: Oxford University Press.

Deliege, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception*, (*4*), 325–360.

Dennett, D.C. (1978b). Skinner skinned. In *Brainstorms* (1981). Montgomer, VT: Bradford Books.

______. (1981). *Brainstorms*: *Philosophical essays on mind and psychology*. Cambridge, Massachusetts: MIT Press.

______. (1991). *Consciousness explained*. London: Penguin Books.

______. (1995). *Darwin's dangerous idea*: *Evolution and the meanings of life*. London: Penguin Books.

______. (2013). *Intuition pumps and other tools for thinking*. London: Penguin Books.

Descartes, R. (1637/1978). *A discourse on method*. London: J. M. Dent & Sons.

Deutsch, D. (1982). *The psychology of music*. New York: Academic Press.

Elias, S. (2012). *Origins of human innovation and creativity*. Amsterdam: Elsevier.

Ellis, G. G. R. (2005). Physics, complexity, and causality. *Nature*, (*435*), 743.

______. (2008). On the nature of causation in complex systems. RSA Conference Paper.

Everett, D. (2005). Cultural constraints on grammar and cognition in pirahã: Another look at the design features of human language. *Current Anthropology*, (*46*)4, 621–646.

Fishman, Y.I. and Steinschneider, M. (2010). Formation of auditory streams. In A. Rees, and A.R. Palmer (Eds.), *The Oxford handbook of auditory science*: *The auditory brain* (pp. 215–245). New York: Oxford University Press.

Forte, A. (1973). *The structure of atonal music*. Yale: Yale University Press.

Gibson, E. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

Gjerdingen, R.O. (1988). *A classic turn of phrase*. Philadelphia: University of Pennsylvania Press.

______. (1991). Defining a prototypical utterance. *Psychomusicology*, (*10*)2, 127–139.

______. (1994). Apparent motion in music?. *Music Perception*, (*11*), 335–370.

______. (1996). Courtly behaviors. *Music Perception*, (*13*), 365–370.

______. (2007). *Music in the galant style*. Oxford: Oxford University Press.

______. (2014) "Historically informed" corpus studies. *Music Perception*, (*31*)3, 192–204.

Gregory, R.L. (1981). *Mind in science*: *A history of explanations in psychology and physics*. Cambridge: Cambridge University Press.

Goehr, L. (1992). *The imaginary museum of musical works*: *An essay in the philosophy of music*. Oxford: Oxford University Press.

Hamanaka, M., Hirata, K. & Tojo, S. (2005). ATTA: Automatic time-span tree analyzer based on extended GTTM. *Proceedings of the Sixth International Conference on Music Information Retrieval*, *ISMIR*, 358–365.

______. (2006). Implementing ‘a generative theory of tonal music’. *Journal of New Music Research*, (*35*), 249–277.

______. (2007). FATTA: Full automatic time-span tree analyzer. *Proceedings of the International Computer Music Conference*, *ICMC*, 153–156.

Hamanaka, M., & Tojo, S. (2009). *Interactive GTTM analyzer*. *Proceedings of the 10th International Society for Music Information Retrieval Conference*, ISMIR, 291–296.

Hansen, N. (2010). The legacy of Lerdahl and Jackendoff’s *A Generative Theory of Tonal Music*: Bridging a significant event in the history of music theory and recent developments in cognitive music research. *Danish Yearbook of Musicology*, (*38*), 33–55.

Hantz, E. (1985). A generative theory of tonal music by F. Lerdahl & R. Jackendoff. Review. *Music Theory Spectrum*, (*7*), 190–202.

Harnad, S. (1990). The symbol grounding problem. *Physica*, (*D42*), 335–346.

Harris, M. (1979). *Cultural materialism*. New York: Random House.

______. (2007). Cultural materialism and behavior analysis: Common problems and radical solutions. *The Behavior Analyst*, (*30*)1, 37–47.

Harvey, D. (1985). A generative theory of tonal music by F. Lerdahl & R. Jackendoff. Review. *Music Analysis*, (*4*)3, 292–303.

Hasty, C.F. (1997). *Meter as rhythm*. Oxford: Oxford University Press.

Hauser, M.D., Chomsky, N., Fitch, W.T. (2002). The faculty of language: What is it? Who has it? And how did it evolve?. *Science*, (*298*), 1569–1579.

Helmholtz, H.L.F. (1863/1875). *On the sensations of tones as a physiological basis for a theory of music*. Longmans, Green, and Co.

Holtmeier, L. (2011). Music in the galant style by Robert Gjerdingen. *Eighteenth-Century Music*. Review. (*8*)2, 307–348.

Holtzmann, S.R. (1977). A program for key determination. *Interface*, (*6*), 29–56.

Humboldt, v.W. (1836/1999). *On language*: *On the diversity of human language construction and its influence on the mental development of the human species*. Cambridge: Cambridge University Press.

Hume, D. (1739/1985). *A treatise of human nature*. London: Penguin Classics.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, (*19*)1, 1–64.

______. (2006). *Sweet anticipation*: *Music and the psychology of expectation*. Cambridge, MA: MIT Press.

______. (2011). *What is a musical work*? *And other curiosities of musical memory*. Exploring the Mind Through Music Conference at Rice University, Houston.

Huron, D. and Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*, (*12*), 154–171.

Ito, J.P. (2014). Koch's metrical theory and Mozart's music: a corpus study. *Music Perception,* (*31*)3, 205–222.

Jan, S. (2000). Replicating sonorities: Towards a memetics of music. *Journal of Memetics – Evolutionary Models of Information Transmission*, (*4*).

______. (2007). *The memetics of music*: *A neo-Darwinian view of musical structure and culture*. Aldershot: Ashgate Publishing Limited.

______. (2010). Memesatz contra *Ursatz*: Memetic perspectives on the aetiology and evolution of musical structure. *Musicae Scientiae*, (*14*)1, 3–50 .

______. (2013). Using galant schemata as evidence for Universal Darwinism. *Interdisciplinary Science Reviews*, (*38*)2, 49–68.

Kaiser, U. (2007a). *Die Notenbücher der Mozarts*: *Als Grundlage der Analyse von W. A. Mozarts Kompositionen 1761–1767*. Kassel: Bärenreiter-Verlag.

______. (2007b). Was ist ein Musikalisches Modell?. *Zeitschrift der Gesellschaft für Musiktheorie*, (*4*)3, 275–289.

Kay, P. and McDaniel, C. (1978). The linguistic significance of the meanings of basic color terms. *Language*, (*54*)3, 610–646.

Kassler, M. (1977). Explication of the middleground of Schenker's theory of tonality. *Miscellanea Musicologica*: *Adelaide Studies in Musicology*, (*9*), 72–81.

Katz, A. (1945). *Challenge to musical tradition*: *A new concept of tonality*. London: Putnam & Co.

Keller, H. (1955). Strict serial technique in Classical music. *Tempo*, 12–16 and 21–24.

Kerman, J. (1985). *Musicology*. London: Fontana.

Klavans, J.L. and Resnik, P. (Eds.) (1997). *The balancing act*: *Combining symbolic and statistical approaches to language*. Cambridge: The MIT Press.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt Brace.

Komar, A.J. (1971). *Theory of suspensions*: *A study of metrical and pitch relations in tonal music*. Princeton: Princeton University Press.

Kramer, J.D. (1988). *The time of music*: *New meanings, new temporalities, new listening strategies*. New York: Schirmer.

Krebs, H. (1999). *Fantasy pieces*: *Metrical dissonance in the music of Robert Schumann*. Oxford: Oxford University Press.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.

______. (1998a). Perceived triad distance: Evidence supporting the psychological reality of neo-Riemannian transformations. *Journal of Music Theory*, (*42*), 265–281.

Krumhansl, C. & Kessler, E.J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology*: *Human Perception and Performance*, (*8*), 24–36.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt Brace.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: The University of Chicago Press.

Lakoff, G. and Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.

Lamal, P.A. (1997). A behavioral view of the visual arts. In P. A. Lamal (Ed.), *Cultural contingencies*: *Behaviour analytic perspectives on cultural practices*. Westport, CT: Praeger Publishing.

Lee, C. 1991. The perception of metrical structure: Experimental evidence and a model. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical Structure* (pp. 59–127). London: Academic Press.

Leman, M. (1995). *Music and schema theory*. Berlin: Springer.

Lerdahl, F. (1988a). Cognitive constraints on compositional systems. In J. Sloboda, (Ed.), *Generative processes in music* (pp. 97–121). Oxford: Oxford University Press. (Reprinted in *Contemporary Music Review*, (*6*), 1992).

______. (1991). Underlying musical schemata. In I. Cross and P. Howell (Eds.), *Representing Musical Structure* (pp. 273–290). New York: Academic Press.

______. (2001). *Tonal pitch space*. Oxford: Oxford University Press.

______. (2009). Genesis and architecture of the GTTM project. *Music Perception*, (*26*)3, 187–194.

______. (2011). Linguistic and musical syntax. Exploring the Mind Through Music Conference, Rice University, Houston.

______. (2012). Musical syntax and its relation to linguistic syntax. Annual LINGUAE lectures at College de France, Paris.

Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: The MIT Press.

Lerdahl, F. and Krumhansl, C.L. (2007). Modelling tonal tension. *Music Perception*, (*24*)4, 329–366.

Lester, J. (1986). *The rhythms of tonal music*. Carbondale and Edwardsville: Southern Illinois University Press.

______. (1990). A classic turn of phrase: Music and the psychology of convention by Robert Gjerdingen. Review. Journal of Music Theory, (34)2, 369–378.

______. (1992). *Compositional theory in the eighteenth century*. Cambridge, Massachusetts: Harvard University Press.

Lewin, D. (1982). A formal theory of generalized tonal functions. *Journal of Music Theory*, (*26*), 23–60.

______. (1987). *Generalized musical intervals and transformations*. New Haven: Yale University Press.

Little M. & Jenne N., (1991). *Dance and the music of J.S. Bach*: *Expanded edition*. Bloomington: Indiana University Press.

Locke, J. (1689/1998). *An essay concerning human understanding*. London: J. M. Dent.

London, J. (1999). Hasty's dichotomy. *Music Theory Spectrum*, (*21*)2, 260–274.

______. (2004). *Hearing in time*: *Psychological aspects of musical meter*. Oxford: Oxford University Press.

Longuet-Higgins, H.C. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, (*1*)4, 424–441.

Longuet-Higgins, H.C. and Lee, C. (1982). The perception of musical rhythms. *Perception*, (*11*), 115–128.

Longuet-Higgins, H.C. and Steedman, M.J. (1971). On interpreting Bach. *Machine Intelligence*, (*6*), 221–241.

Ludvig, E.A. (2003). Why Pinker needs behaviorism: A critique of *The Blank Slate*, *by Steven Pinker*. *Behavior and Philosophy*, (*31*), 139–143.

Lyotard, J-F. (1979a). *The postmodern condition*: *A report on knowledge*, tr. G. Bennington and B. Massumi, Manchester: Manchester University Press (1984).

Mandler, G. (1984). *Stories*, *scripts*, *and scenes*: *Aspects of schema theory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Marsden, A. (2001). Representing melodic patterns as networks of elaborations. *Computers and the Humanities*, (*35*), 37–54.

______. (2005). Generative structural representation of tonal music. *Journal of New Music Research*, (*34*), 409–428.

______. (2007). Automatic derivation of musical structure: A tool for research on Schenkerian analysis. *Proceedings of the International Conference on Music Information Retrieval*, *ISMIR*, 55–58.

______. (2010). Schenkerian analysis by computer: A proof of concept. *Journal of New Musical Research*, (*39*)3, 269–289.

______. (2012). Interrogating melodic similarity: A definitive phenomenon or the product of interpretation?. *Journal of New Musical Research,* (*41*)4, 323–335.

______. (2013). Towards computable procedures for deriving tree structures in music: Context dependency in GTTM and Schenkerian theory. *Proceedings of the Sound and Music Computing Conference*, 360–367.

Martin, P., and Bateson, P. (2007). *Measuring behaviour*: *An introductory guide*: *Third edition*. Cambridge: Cambridge University Press.

Maxwell, H.J. (1992). An expert system for harmonic analysis of tonal music. In M. Balaban, K. Ebcioglu, & O. Laske (Eds.), *Understanding Music with AI* (pp. 335–53), Cambridge, MA: MIT Press.

McCarthy, J. and Hayes, P. 1969. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie (Eds.), *Machine Intelligence 4*, (pp. 463–502).

McFarland, M. (2012). Bill Evans and the limits of Schenkerian analysis. *Journal of Schenkerian Studies*, (*6*).

McKee, E. (2004). Extended anacruses in Mozart's instrumental music. *Theory and Practice*, (*29*), 1–37.

Meyer, L.B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

______. (1973). *Explaining music*. Berkeley: University of California Press.

______. (1989). *Style and music*: *Theory*, *history*, *and ideology*. Philadelphia: University of Pennsylvania Press.

Meyer, L.B. & Rosner, B.S. (1980). Exploiting limits: Creation, archetypes, and style change. *Daedalus*, (*109*)2, 177–205.

Miranda, E.R., Kirby, S., & Todd, P.M. (2003). On computational models of the evolution of music: From the origins of musical taste to the emergence of grammars. *Contemporary Music Review*, (*22*)2, 91–110.

Minsky, M. (1986). *The society of mind*. New York: Touchstone.

Mirka, D. (2009). *Metric manipulations in Haydn and Mozart*: *Chamber music for strings*: 1787–1791. New York: Oxford University Press.

Mithen, S. (2005). *The singing Neanderthal*: *The origin of music, language*, *mind and body*. London: Weidenfeld & Nicolson.

Narmour, E. (1977). *Beyond Schenkerism*: *Towards an implication-realization model*. Chicago: The University of Chicago Press.

______. (1990). *The analysis and cognition of basic melodic structures*: *The implication–realization model*. Chicago: The University of Chicago Press.

Nattiez, J-J. (1990). *Music and discourse*: *Toward a semiology of music*, tr. C. Abbate, Princeton: Princeton University Press.

Neuwirth, M. (2008). Music in the galant style, by R.O. Gjerdingen. Review. *Zeitschrift der Gesellschaft für Musiktheorie*, (5)2–3, 401–409.

Nietsche, F. (1886/2002). *Beyond good and evil*: *Prelude to a philosophy of the future*. Cambridge: Cambridge University Press.

Noorden, L.P.A.S.v. (1975). *Temporal coherence in the perception of tone sequences*. Unpublished Ph.D. diss., Eindhoven University of Technology.

Okasha, S. (2012). Emergence, hierarchy, and top-down causation in evolutionary biology. *Interface Focus*, (*2*), 49–54.

Okasha, S. (2006). *Evolution and the levels of selection*. Oxford: Oxford University Press.

Ottman, R. (1986). *Music for sight singing*. Englewood Cliffs, NJ: Prentice-Hall.

Pagel, M. (2012). *Wired for culture*. Great Britain: Allen Lane.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, (*11*), 409–464.

Pavlov, I.P. (1927). *Conditioned reflexes*: *An investigation of the physiological activity of the cerebral cortex*, tr. G. V. Anrep. London: Oxford University Press.

Pearce, M.T. and Wiggins, G.A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*, (*23*)5, 377–405.

______. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, (*4*)4, 625–652.

Pearl, J. (2000). *Causality*. Cambridge: Cambridge University Press.

Peel, J. & Slawson, W. (1984). A generative theory of tonal music, by F. Lerdahl & R. Jackendoff. Review. *Journal of Music Theory*, (28)2, 271–294.

Piaget, J. (1952). *The origin of intelligence in children*. New York: International University Press, Inc.

Pinker, S. (1997). *How the mind works*. New York: W.W. Norton & Company.

______. (2002). *The blank slate*: *The modern denial of human nature*. New York: Viking Press.

Pinker, S. & Jackendoff, R. (2005). The faculty of language: What's special about it. *Cognition*, (*95*), 201–236.

Piston, W. (1941). *Harmony*. New York: W.W. Norton & Company, Inc.

Pople, A. (2004). Modeling musical structure. In E. Clarke, and N. Cook (Eds.), *Empirical musicology*: *Aims*, *methods*, *prospects* (pp. 127–156). Oxford: Oxford University Press.

Popper, K. (1935/2002). *The logic of scientific discovery*. London: Routledge.

Popper, K. and Eccles, J. (1977). *The self and its brain*. Berlin, London: Springer-Verlag.

Povel, D. –J. & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, (*2*), 411–440.

Prince, J.B. and Schmuckler, M.A. (2014). The tonal-metric hierarchy. *Music Perception*, (*31*)3, 254–270.

Rameau, J-P. (1722/1971). *Treatise on harmony*. United States: Dover Publications, Inc.

Ratner, L.G. (1956). *Classic music*: *Expression*, *form*, *and style*. New York: Schirmer.

Reti, R. (1951). *The thematic process in music*. New York: Macmillan.

Riemann, H. (1893). Vereinfachte Harmonielehre oder die Lehre von Tonalen Funktionen der Akkorde. London: Augener.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, (*5*)1, 35–53.

Rohrmeier, M. and Neuwirth, M. (2015). Towards a syntax of the classical cadence. In Neuwirth, M. and Berge, P. (Eds.), *What is a Cadence*: *Theoretical and analytical perspectives on cadences in the Classical repertoire* (pp. 287–338). Leuven: Leuven University Press.

Rosch, E. (1973). (Eleanor Heider) Natural categories. *Cognitive Psychology*, (*4*), 328–350.

______. (1975a). Cognitive reference points. *Cognitive Psychology*, 532–547.

______. (1975b). Cognitive representations of semantic categories. *Journal of Experimental Psychology*, (*104*), 192–233.

______. 1978. Principles of categorization. In E. Rosch, and B. Lloyd (Eds.) (1978), 27–48.

Rosch, E. & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, (*7*), 573–605.

Rosch, E., Mervis, C., Gray, Johnson, D., Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, (*8*), 382–439.

Rosen, C. (1971). *The Classical style*. New York: W.W. Norton & Company.

Rosen, C. (1980). *Sonata forms*. New York: W.W. Norton & Company.

Rosenthal, D. (1992). Emulation of human rhythm perception. *Computer Music Journal*, (*16*)1, 64–76.

Rothstein, W. (1989). *Phrase rhythm in tonal music*. New York: Schirmer Books.

______. (1995). Beethoven with and without kunstgepräng: Metrical ambiguity reconsidered. In L. Lockwood and J. Webster (Eds.), *Beethoven Forum* (pp. 4165–4193). Lincoln: University of Nebraska Press.

Rumelhart, D.E. (1975). Notes on a schema for stories. In D. Norman and D.E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: Freeman.

______. (1980). Schemata: the building blocks of cognition. In R.J. Spiro et al. (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Lawrence Erlbaum.

Rummelhart, D.E. and Ortony, A. (1977). The representations of knowledge in memory. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Russell, B. (1946). *History of Western philosophy*. London: George Allen & Unwin Ltd.

Ryle, G. (1949/2009). *The concept of mind*. Abingdon: Routledge

Salzer, F. (1952). *Structural hearing*: *Tonal coherence in music*. New York: Dover.

Salzer, F. and Schachter, C. (1969). *Counterpoint in composition*. New York: Columbia University Press.

Sanguinetti, G. (2012). *The art of partimenti*: *History*, *theory*, *and practice*. Oxford: Oxford University Press.

Schank, R. C, and Abelson, R.P. (1977). *Scripts*, *plans*, *goals and understanding*: *An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schachter, C. (1976). Rhythm and linear analysis: A preliminary study. In F. Salzer (Ed.), *The Music Forum*, Vol. 4 (pp. 281–334). New York: Columbia University Press.

______. (1980). Rhythm and linear analysis: Durational reduction. In F. Salzer and C. Schachter (Eds.), *The music forum*, Vol. 5 (pp. 197–232). New York: Columbia University Press.

______. (1987). Rhythm and linear analysis: some aspects of meter. In F. Salzer and C. Schachter (Eds.), *The music forum*, Vol. 6, Part 1 (pp. 1–59). New York: Columbia University Press.

______. (1999). *Unfoldings*: *Essays in Schenkerian theory and analysis*. Oxford: Oxford University Press.

Schönberg, A. (1911/1978). *Theory of harmony*, tr. T.R. Carter. Berkeley: University of California Press.

______. (1954). *Structural functions of harmony*. London: Williams and Norgate Ltd.

______. 1967. *Fundamentals of musical composition*. London: Faber and Faber Ltd.

Schenker, H. (1935/1979). *Free composition*, tr. E. Oster. New York: Longman.

Searle, J.R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences* (*3*)3, 417–457.

Skinner, B.F. (1953). *Science and human behavior*. USA: The Free Press.

______. (1957). *Verbal behavior*. USA: Appleton-Century-Crofts.

______. (1981). Selection by consequences. *Science*, (*213*), 501–504.

Sloboda, J.A. (1985). *The musical mind*. Oxford: Clarendon Press.

______. (1987). *Generative processes in music*. Oxford: Clarendon Press.

Smoliar, S.W. (1980). A computer aid for Schenkerian analysis. *Computer Music Journal*, (*4*), 41–59.

Snyder, B. (2000). *Music and memory*: *An introduction*. Cambridge, MA: MIT Press.

Steedman, (1977). The perception of musical rhythm and meter. *Perception*, (*6*), 555–570.

Suttcliffe, W.D. (2009). Music in the galant style by R.O. Gjerdingen. Review. *Music & Letters*, (90)3, 468–473.

Swain, J.P. (2002). *Harmonic rhythm*: *Analysis and interpretation*. Oxford: Oxford University Press.

Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, Mass: MIT Press.

______. (1995). Motivic perception and modularity. *Music Perception*, (*13*), 141–169.

______. (2006). Music in the Galant Style by R.O. Gjerdingen. Review. *Music & Letters*, (50)2, 277–290.

______. (2007). *Music and probability*. Cambridge, Mass: MIT Press.

Tenney, J. & Polansky, L. (1980). Temporal gestalt perception in music. *Journal of Music Theory*, (*24*), 205–241.

Wertheimer, M. (1923). Untersuchung zur Lehre von der Gestalt. II. *Psychologische Forschung*, (4), 301–350.

West, W., Howell, P., & Cross, I., (1991). Musical structure and knowledge representation. In I. Cross and P. Howell (Eds.), *Representing musical structure* (pp. 1–30). New York: Academic Press.

Wiggins, G. (2008). A review of computer models of musical creativity by David Cope. *Literary and Linguistic Computing*, (*23*)1, 109–116.

Wiggins, G.A., Müllensiefen, D., Pierce, M.T. (2010). On the non-existence of music: Why music theory is a figment of the imagination. *Musicae Scientiae*, *Discussion Forum*, 5, 231–255.

Winograd, T. (1968). Linguistics and the computer analysis of tonal harmony. *Journal of Music Theory*, (*12*), 2–49.

Yeston, M. (1979). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.

Zatorre, R. (2012). Music in the brain: Pitch, imagery, and emotion. Exploring the Mind Through Music Conference. Rice University, Houston.

Zbikowski, L. (2002). *Conceptualizing music*: *Cognitive structure*, *theory*, *and analysis*. New York: Oxford University Press.

Zuckerkandl, V. (1956). *Sound and symbol*: *Music and the external world*. Princeton: Princeton University Press.