System & Contrast : a Polymorphous Model of the Inner Organization of Structural Segments within Music Pieces (Original Extensive Version)

Frédéric Bimbot, Emmanuel Deruty, Gabriel Sargent, Emmanuel Vincent

Publications Internes de l'IRISA
ISSN : 2102-6327
PI 1999 – décembre 2012

SYSTEM & CONTRAST :

A POLYMORPHOUS MODEL OF THE INNER ORGANIZATION OF STRUCTURAL SEGMENTS WITHIN MUSIC PIECES

(ORIGINAL EXTENSIVE VERSION)

Frédéric BIMBOT [1], Emmanuel DERUTY [2], Gabriel SARGENT [3], Emmanuel VINCENT [4]

METISS Research Group, IRISA - Dept. DSIR,
[1]CNRS - UMR 6074, [2]INRIA - Rennes Bretagne Atlantique, [3]Université Rennes 1
Campus Universitaire de Beaulieu, 35042 Rennes cedex, France

frederic.bimbot@irisa.fr emmanuel.deruty@gmail.com gabriel.sargent@irisa.fr emmanuel.vincent@inria.fr

Abstract

At a large timescale, music pieces can be described as the succession of structural segments which form the global organization of the piece.

The present article proposes a model called "System & Contrast", which aims at describing the *inner organization* of such structural segments in terms of : (i) a *carrier system*, i.e. a sequence of morphological elements forming a matrix network of self-deducible syntagmatic relationships and (ii) a *contrast*, i.e. a substitutive element, usually the last one, which partly departs from the logic of the system.

The S&C model applies at several timescales and to a wide variety of musical dimensions in a very polymorphous way, therefore offering an efficient meta-description of mid-level musical content.

Key-words : music structure, form, semiotics, semiology, relational graph, music analysis, music signal processing, music information retrieval, musicology, system & contrast

Résumé

A une échelle macroscopique, les morceaux de musique peuvent être décrits comme une suite de segments structurels qui participent à l'organisation globale du morceau.

Cet article propose un modèle dénommé "Système-Contraste", qui vise à décrire l'*organisation interne* des segments structurels en tant que (i) un *système porteur*, c'est-à-dire une séquence d'éléments morphologiques formant un réseau matriciel de relations syntagmatiques auto-déductibles et (ii) un *contraste*, c'est-à-dire un élément de substitution, situé généralement en dernière position, et qui s'écarte partiellement de la logique du système porteur.

Le modèle Système-Contraste peut s'appliquer de façon très polymorphe, à différentes échelles de temps et à un vaste ensemble de dimensions musicales, offrant ainsi une puissante méta-description du contenu musical aux échelles intermédiaires.

Mots-clefs : structure musicale, forme, sémiotique, sémiologie, graphe relationnel, analyse musicale, traitement du signal musical, recherche d'information musicale, musicologie, système-contraste

Contents

1. INTRODUCTION

1.1 Context and focus

At small timescales (i.e. typically up to 1 second) a music piece is usually described as a combination of unitary elements such as pitches, durations or chords, drawn from a limited inventory of conventionally pre-defined items. However, above a certain timescale, the content of the piece can furthermore be described in terms of piece-specific objects whose layout contributes to the structure of musical segments at increasing scales (cells, motifs, phrases, sections, …) and ultimately, to the global organization of the music piece (often called its *form*).

Describing unequivocally the organization of music pieces in terms of structural elements turns out to be an open problem. Several recent studies in the context of Music Information Retrieval (MIR) have been aiming at characterizing structural units and form in music pieces, so as to produce consistent annotated resources for research (Peeters & Deruty, 2009 ; Bimbot, Le Blouch, Sargent & Vincent, 2010 ; Bimbot, Deruty, Sargent & Vincent, 2011 ; Smith, Burgoyne, Fujinaga, De Roure, Downie, 2011). However, they all have been facing difficulties in formulating general properties and criteria which could qualify objectively and unambiguously the structural units, independently of the music genre, style or function and without giving priority to a specific musical dimension (melody, harmony, drums, …) over the others.

This situation reflects a gap between, on the one hand, the profuse literature dedicated to traditional analysis of music structure and form in musicology (Perone 1998), which tends to account for many different aspects of the question, and, on the other hand, the need for generic schematic concepts focused on (and suited to) the production of standardized resources, which requires some convergence towards a homogeneous framework. Gradually bridging this gap may indeed benefit both to the development of efficient algorithms for automatic music processing in MIR and to a more global understanding of the underlying properties of music structure.

Over the past few years, our research group has been investigating the issue of music structure description, both from fundamental and experimental viewpoints. The experience acquired through the annotation, discussion and adjudication of several hundreds of pieces has gradually led us to develop a number of concepts and procedures for producing consistent representations of music structure. Our methodology has been addressing the entire range of tasks involved in music structure description, namely segmentation, characterization and labeling of structural units, with the concern that this methodology should be applicable to a wide variety of music pieces (Bimbot, Deruty, Sargent & Vincent, 2012).

In this article, we focus on one particular aspect of this methodology : how to formulate the properties which characterize structural segments, i.e. sections of the musical content at an *intermediate* timescale (typically 10-20 s) bound to constitute relevant units to describe the structure of a music piece at a *large* timescale (namely, its entire span).

Our approach relies on a fundamental pre-requisite : a consistent definition of structural segments should not primarily be based on particular properties of their intrinsic substance, as those properties are bound to be "dangerously" idiosyncratic (i.e. too specific to a given music genre or to a particular musical dimension). An alternative, which is elaborated on in the present article, is to define structural segments primarily on the basis of their *inner organization*. For this purpose, we introduce a new approach, the System & Contrast (S&C) model, which aims at providing a coherent and comprehensive way to describe internal patterns forming the backbone of structural segments.

1.2 Positioning and objectives

The main entry for *structure* in the Oxford dictionary yields : "*the arrangement of and relations between the parts or elements of something complex*". In line with this view of structure, the objective of this article is to define a framework which is able to account for the formal arrangement and relations observed within musical segments in a wide variety of situations and as independently as possible of the nature of their actual content.

For that reason, rather than trying to characterize structural units by an expert set of properties, rules or cues defined on a multitude of musical dimensions and conditioned to a particular musical genre, we approach the determination of structural segments as a *model matching* problem. This conception relates to prototype theory addressed in (Deliège, 2001) and particularly to the notion of "abstracted central tendency", as mentioned in (Lamont, 2001). In our work, the prototypic model is designed and formulated so as to be applicable in a versatile way to several musical dimensions and to a wide variety of music pieces. As we will develop it in this article, our model can also be approached with concepts resorting to Information and Communication Sciences.

The basic idea behind the S&C model is that a structural segment draws its consistency from the fact that it constitutes a *system*, namely *"an interdependent group of items forming a unified whole"* (definition of the Merriam-Webster dictionary), or, in other words, *"an entity of internal dependencies"* (Hjelmslev, 1959). Distinct segments result from distinct systems, but all such systems are assumed to be governed by a common model, which therefore needs to be *polymorphous*, in the sense that it is able to accommodate a wide variety of forms, types and realizations of structural segments within its framework.

Similarly to the neutral level analysis suggested in (Nattiez, 1987) the S&C approach does not seek to decipher and uncover the *message* behind a musical segment. It only aims at describing its *organization* with reference to an archetypal model. Following Ruwet's conception (1966), we approach music as a semiotic process, on which we focus our interest on the *structure* of the *code*. This may result in a level of description of the musical message that does not correspond to that intended by the composer and that does not account for the actual musical language used by the composer (Nattiez, 1987, p. 43). Neither may it necessarily reflect perceptual characteristics that listeners would primarily identify as the most salient ones in the musical narration.

In this respect, the S&C model does not suggest any kind of musical "truth". It is neither prescriptive, nor interpretative. It merely proposes a standardized description of organizational patterns which take place within structural segments. It thus provides a generic framework to which the observed musical content can be matched against and used to arbitrate between multiple description hypotheses, by means of objective criteria, formulated in reference to this model.

In accordance with the principles on which is grounded the work of Lerdahl and Jackendoff (1983), we consider that the determination of structural segments relies on a grouping process, based on the relationships between inner constituents of the segment. However, rather than assuming a hierarchical model, our approach rests on a *matrix* structure. This relaxes the adjacency constraint which governs the conception developed in "*a generative theory of tonal music*", strictly based on relations between contiguous segments.

The S&C model rests on premises which closely relate to those of the Implication-Realization model proposed by Narmour (1990, 1992)[1], originally for melodic structures, later extended towards the concept of cognitive *rule-mapping* (Narmour 2000) and whose principles are considered by its author to be applicable to other musical dimensions (Narmour 2000, pp. 365-372). Indeed, like ours, Narmour's approach aims at accounting for aspects which govern cognitive expectation schemes in the musical discourse, and the two models undoubtedly share common views on the importance of realization, denial or surprise in the determination of structural units in music.

Proposing a new model (and, implicitly, a new underlying "theory") raises the question on how to assess its soundness. In the present case, a tractable quantitative evaluation of the model turns out to be difficult to conceive or at least quite premature. Therefore, we support the claims developed in this article by (i) a *body of evidence* showing that the proposed model is able to relevantly account for a wide variety of structural phenomena within a significant domain of validity, (ii) *consistency proofs* in the sense that the proposed model does not contain internal contradictions and (iii) *backward compatibility*, i.e. the ability of the proposed framework to account for well-known structural schemes.

1.3 Scope and working hypotheses

Before introducing the S&C model, it is important to formulate explicitly a number of working hypotheses and related considerations, so as to clarify the precise scope of this work.

- The proposed model aims at covering a wide range of music categories, genres and styles. We do not claim that the model is able to handle absolutely all types of music, but, based on our own experience, it deals relatively well with music from the late baroque, classical, romantic eras ; it is particularly well suited to describe pop, rock, dance, jazz, blues, urban music, techno, occidental folk, etc… ; but it can fail to account for structures found in modern and 20th century music, as well as in many types of contemporary music, whose constructions can depart radically from "conventional" practice. This remark is not considered as a matter of disqualification of the model but rather as an encouragement to use it for a large range of music genres, though with some discernment.
- In the same way, we do not claim that the S&C model can cover all aspects of form, in the many dimensions along which it can be approached, analyzed and explained, in particular, structural "functions", in the sense of Spencer & Temko (1988). As developed further in this article, the S&C is intended to account for what we call the *morpho-syntagmatic* structure of musical segments. This term is chosen in analogy with structural linguistics (Chomsky, 1957), where it refers to a purely grammatical approach of the message based on the form and relations of its constitutive elements, without regard to functional (or semantic) considerations. One strength of the S&C lies in its *polymorphism*, i.e. its capacity to encompass under a single formalism a wide variety of structural arrangements, at several scales, along multiple musical dimensions and independently of their role or value in the musical narration. Therefore, the scope of the S&C model is strictly that of a *descriptive model* of organizational aspects of the structure of music, which can feed and fuel compositional, perceptual or functional analyses of a musical passage. It certainly does not replace such analyses, but it offers them a framework on which they may rest and from which they can benefit.
- In its widest sense, music analysis can be defined as the process of examining musical material so as to determine a preferred "explanation" (i.e. a most plausible description) of the musical content within a given framework. In that sense, the S&C model should be viewed as an analysis *tool*. As in any analysis process, S&C modeling involves breaking the music passage down into smaller parts, quantizing them into discrete states, and simplify their representations along some dimensions (including the timescale). These "reductions" turn out to be necessary to simplify the characterization of the objects which are taking part to the definition of the system. In fact, the proposed approach is itself *reductionist*, in the sense that it aims at fitting as much as possible the observed data within the mod-

[1] See also (Cumming, 1992).

el. The resulting approximations should not be viewed as weakening the principles of the approach : they are indeed necessary to get rid of surface effects, as long, of course, as they do not appear to radically alter the relevance of the analysis, by distorting some essential factor.

- A noteworthy quality of the proposed approach is to be based essentially on *deductive* reasoning, i.e. using only information available in the piece itself. This viewpoint differs from inductive reasoning which consists in inferring a solution in reference to similar configurations observed in other situations. Behind the concept of deductive reasoning lies the hypothesis that information compression processes play an essential role in the cognitive emergence of formal schemes, as thoroughly developed in (Levy, 2003) within the scope of music analysis and composition. Of course, deductive reasoning may turn out to be even easier when the analyst is familiar with the type of music under consideration. But altogether, the proposed approach has this interesting property that, with a little training, it can be dealt with by "experienced listeners" (Lerdahl & Jackendoff, 1983) without requiring them to be instructed expert musicologists.

1.4 Overview of the article

In order to introduce the actual matter of this article, we have chosen to start by presenting (in section 2) the principles of the S&C model in a completely intuitive and metaphoric manner, so that the reader can immediately grasp the basic notions behind the proposed approach. Then, section 3 formalizes the S&C model in its standard form (square system), and introduces central concepts such as those of "carrier system", "contrast", "morphological elements" and "syntagmatic functions".

It is only in section 4 that the S&C model is connected to the domain of music, by making explicit links between the model components and actual musical dimensions, properties, elements, relations and patterns occurring in musical segments. In particular, this section shows how the S&C model can encompass several well-known structural configuration into a single framework.

Section 5 extends the basic (square) S&C model to a well-definable set of S&C typologies (in particular pentadic and hexadic systems), which are able to account for a wide variety of sophisticated constructions based on a compact network of relationships between subdivisions of the structural segment. This leads to the codification of S&Cs into a limited inventory of "hypermetric shapes".

Section 6 turns towards information theory and relates, though the MDL (Minimum Description Length) criterion, the proposed analysis paradigm to a compression scheme which aims at finding the simplest way to explain a musical segment given the S&C as a class of models. We detail how the S&C analysis of a passage can be obtained as the result of a model matching process which arbitrates between several hypothesis by optimizing the overall economy of the resulting description.

In section 7, we connect the S&C model to a number of basic concepts resorting to information and communication sciences and in particular, we discuss the roles played by the contrast within the S&C model, thus highlighting several of its functions.

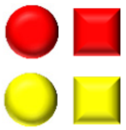
Before concluding this article, section 8 discusses briefly on the potential of the model in various application domains such as music analysis, music modeling, music information retrieval, computer-aided music and music education.

Sixteen examples of musical passages from a wide variety of musical genres accompany this article to illustrate the ability of the S&C model to describe, in a generic way, a multitude of configurations pertaining to the inner structure of musical segments.

Throughout the article, care has been taken to expose the various arguments supporting the proposed approach with the concern of making them as widely accessible as possible (and hopefully relevant) to colleagues from various disciplines interested in music modeling, beyond MIR and engineering sciences. Indeed, a number of concepts which have been guiding this work derive or are inspired from the fields of semiotics and linguistics, and connect to the domains of cognitive science, communication sciences and information theory.

2. INTUITIVE PRESENTATION

2.1 The square system



These 4 elements form a *system* based on a combination of two binary *oppositions*, in terms of shape and color. We will call this system a *square system*.

Figure 1 shows a few examples of square systems, for which it is easy to figure out which are the properties used as oppostions, and therefore to easily *explain* the system[1].

Figure 1 : Four examples of square systems

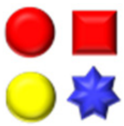
2.2 The contrast

A fundamental property of a square system is its redundancy. Indeed, Figure 2 depicts a few incomplete square systems, i.e. systems for which the 4th element is missing (and replaced by a question mark).

Figure 2 : Four incomplete square systems

As can be easily experienced by the reader, some properties of the 4th element are predictable and can be logically *deduced* from the knowledge of the first 3 elements[2].

As a consequence, it is easy to determine, on the basis of the exposition of 3 elements and the presentation of a fourth one, whether this 4th element matches or deviates from the system, and, if so, <u>in what respect</u>.



These 4 elements now form what we call a *System & Contrast (S&C).* The shape and color properties of the 4th element both contradict the combination expected in 4th position, given the first three elements. The 4th element creates a logical *contrast* <u>within the system</u>.

2.3 Carrier system and contrasting properties

The characterization of a system and its contrast requires the simultaneous determination of the set of properties which form the system and the identification of those which take part in the contrast. Figure 3 illustrates several configurations where the contrastive property varies over the same baseline square system, which we call the *carrier system.*

Carrier system (no contrast) | Colour-based contrast | Shape-based contrast | Size-based contrast

Figure 3 : Examples of various contrasts, based on the same carrier system

As developed later in this article, it is worth noting that the contrastive properties act as a *logical (or digital)* <u>*modulation*</u> of the *information* conveyed by the carrier system.

---

[1] Note however that some properties may not participate to the system, as in the 4th example of figure 1, where the font does not show any systematic behavior.

[2] "Germany", an orange diamond, the digit "8" in dark green and any item with a NW-SE orientation. In fact, the 4th system in figure 2 could be considered as a complete system, since the orientation of the question mark is consistent with that of the rest of the system, and no other property seems to show any systematic behavior.

2.4 Analyzing an S&C

Let's now consider the following two quadruplets of elements :

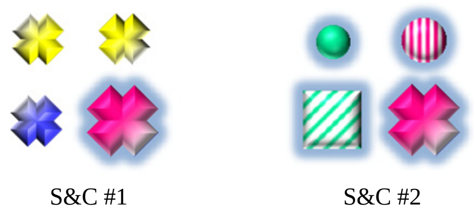


Figure 4 : Examples of two S&Cs (analyzed in the text).

A careful examination of S&C #1 leads to the following analysis : the 5 properties which appear to be relevant to explain the elements in the system are *shape, color, size, halo* and a very subtle *shade orientation*. Given the first three items, the logical element in position 4 should be a *medium-size blue cross* with its pale branch pointing *downwards*, and having *no halo*. The 4^{th} element is indeed a *cross* with its pale branch pointing *downwards* but it is *large, red* and surrounded *with a halo*. In S&C #1, the contrast therefore affects 3 properties : *color, size* and *halo*.

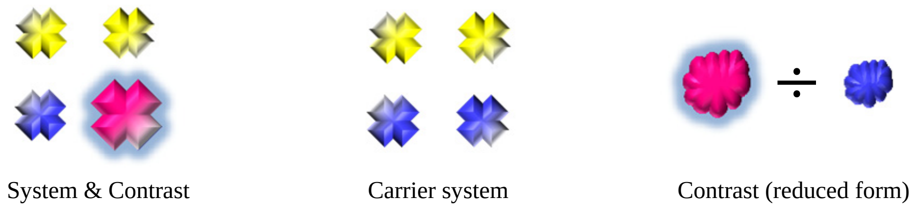


Figure 5 : Result of the analysis of S&C #1

Let's now consider S&C #2 : *shape, color, size, homogeneity, halo, texture* and potentially *orientation* are properties of the elements in the system. However, (i) *texture* varies erratically and can therefore be considered as an off-system property and (ii) the status of *orientation* as a systematic property is not decidable, because it is not possible to evaluate it for the circles. Among the 5 remaining properties, only *shape, size* and *homogeneity* participate to the contrast. Indeed, the 4^{th} element is a *large shaded cross* instead of being a *very-large unshaded square*. Figure 6 summarizes the result of the analysis of S&C #2

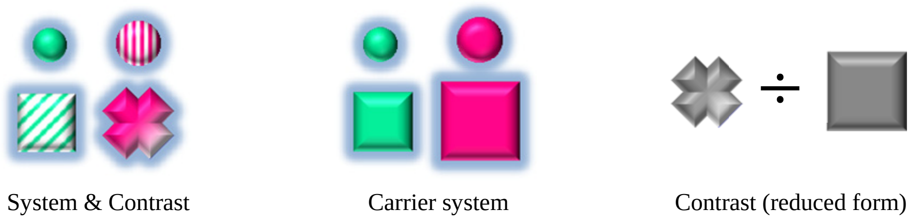


Figure 6 : Result of the analysis of S&C #2

These examples illustrate how a 2 × 2 square S&C forms a set of elements for which the *matricial* combination of several properties contributes to the cohesion of the system. At the same time, the existence of some redundancy in the system offers the possibility to insert some additional (contrastive) information in the 4^{th} position, which can be deduced from the way the properties of individual items vary across the system.

In Figures 5 and 6, the contrast is denoted as $\mathcal{P} \div \hat{\mathcal{P}}$, where $\mathcal{P}$ is the set of observed properties and $\hat{\mathcal{P}}$, the expected ones, *reduced* to the subset of the carrier properties which actually differ in the contrast (as a numerical fraction would be simplified into its irreducible form).

It is worth noting that, in the two examples here, a same element in 4^{th} position presents two very different contrastive *values* within the two systems.

3. FORMALIZATION

3.1 Specification of the carrier system

A square system (in its carrier form) can be denoted as :

$$S_0 = \begin{matrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{matrix}$$

As S_0 forms a square system, a network of *similarity* relations exists between its 4 elements :

- horizontal relation : $x_{01} = f(x_{00})$
- vertical relation : $x_{10} = g(x_{00})$
- diagonal relation : $x_{11} = g\big(f(x_{00})\big) = f\big(g(x_{00})\big)$

This can also be stated as a logical proposition :

> x_{11} is to x_{10} what x_{01} is to x_{00}
> **and**
> x_{11} is to x_{01} what x_{10} is to x_{00}

This is nothing else than the *generalization* of the well-known "rule of three", i.e. the relationship between 4 numbers forming a system of proportions.

Even if many *descriptive* properties are involved in characterizing all the elements in S_0, the relations f and g may apply to only a subset of the properties of the elements of S_0, which constitute the *structuring* properties of the system.

Altogether, the carrier system boils down to an initial element (x_{00}) and a *redundant* network of relations (f, g and $gof = fog$).

3.2 Formulation of the contrast

Following similar notations as in the previous subsection, the System & Contrast can be noted :

$$S = \begin{matrix} x_{00} & x_{01} \\ x_{10} & \bar{x}_{11} \end{matrix}$$

Whereas the horizontal and vertical similarity relations (f, g) remain identical to that of the carrier system, we now have a specific diagonal relation : $\bar{x}_{11} = k(x_{00})$, with $k \neq gof$.

A contrast results from the *disparity* between k and gof and this disparity can itself be viewed as an additional relation γ which expresses the deviation of the actual element $\bar{x}_{11}$ with respect to the (virtual) *expected* one $x_{11} = (gof)(x_{00})$, i.e :

$$\gamma = ko(gof)^{-1}$$

As a result of the presence of the contrast, we now have the following situation :

> $\bar{x}_{11}$ is **not** to x_{10} what x_{01} is to x_{00}
> **and/or**
> $\bar{x}_{11}$ is **not** to x_{01} what x_{10} is to x_{00}

Element $\bar{x}_{11}$ is breaking the "natural" flow of events and creates a logical rupture. Relation γ thus appears as a discordance in the system, which can be detected in reference to the other elements by first deducing and then factoring out the properties of the *carrier* system S_0. The contrast appears as the reduced form of $\bar{x}_{11} \div x_{11}$.

Relation γ may not apply to all structuring properties but only to a subset of them, which constitute the set of *contrastive properties*, this set being possibly empty.

3.3 S&C constituents and terminology

The systems considered in this paper up to now are 2×2 matrices. Therefore, we will call them 2×2 (square) systems, but in a forthcoming section, we consider systems with other *shapes*.

We designate as *morphological elements*, the set of items x_{ij} forming the system S. Among them, the initial element x_{00} is called the *primer* and the last one, the *contrast*. If γ is the identity function, there is no actual contrast at the end of the system, and the S&C is said to be *plain*.

In this work, relations f and g are referred to as *syntagmatic relations* or *syntagmatic functions* : they correspond to position-dependent *relations* between elements within the system S which can equivalently be understood as formal *similarity functions* between two morphological elements or, from a more generativist viewpoint, *transformations* which map a given morphological element to another.

In the most general case, syntagmatic relations are multidimensional, i.e. they apply to several variables / properties at a time. However, we assume that they can be decomposed as a vector of one-dimensional functions operating on each dimension individually. In the simplest case, syntagmatic functions can be the identity function (*id*). Conversely, they may need to be defined *in extenso*, i.e. as a point-to-point mapping specifying how each value of each variable matches each other in the two related elements (cf. function *new* defined in section 4.4).

The quadruplet (x_{00}, f, g, γ) constitutes a *morpho-syntagmatic description* of system S in reference to the S&C model, or in short an *S&C description* of S. This quadruplet can be viewed metaphorically as the "genetic code" of the system.

3.4 Visual representation of the S&C model

Figure 7 below depicts a 2 × 2 S&C unfolded along the timeline (as $x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}$), on which are represented the main constituents of the system : the morphological elements x_{ij}, the syntagmatic relations *f*, *g* and the contrast function γ.

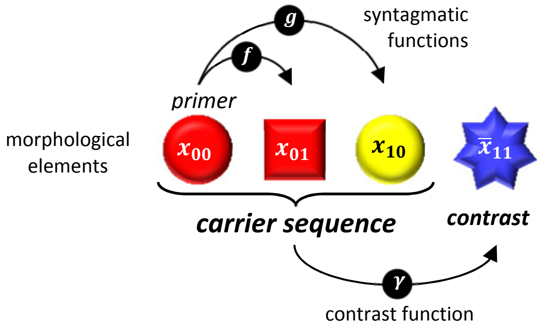


Figure 7 : A schematic view of a 2×2 square S&C model components (unfolded form)

4. THE S&C MODEL FOR MUSIC

In this section we develop how the S&C model can be applied to music, and in particular how it appears to be interestingly well-suited to describe the internal organization of structural segments.

4.1 Delinearization and causality principle

A fundamental assumption which underlies the model proposed in this article is that, even though music appears as a *sequential* presentation of acoustic (or symbolic) material, it is in fact latently governed by *matricial* relationships between its elements. Analyzing a musical content as an S&C therefore implies an operation which can be called *delinearization*.

At the same time, we favor a *causality principle* in the sense that the *direction* of the relationships between elements within the system is constrained to be causal, i.e. in accordance with the order in which these elements occur in the musical flow, and in reference to the first element of the system. In particular, the contrast is always assumed to be in last position, even though, in some cases, it would also be possible to explain the system on the basis of a disparity affecting another element than the last one.

From now on, we will represent square systems in an unfolded way rather than as a 2×2 matrix, but it should be kept in mind that they must be apprehended in their matrix form.

4.2 Musical dimensions and structuring properties

Analogously to the visual examples illustrated in section 2, the structuring properties involved in the description of an S&C can rely on virtually <u>any combination of dimensions</u> of the musical content, provided they evolve in an *organized* manner within a musical segment and form a consistent set of relatively simple relationships.

For instance, a non-exhaustive list of possible structuring properties is :

- Melody contour / melodic intervals / support notes of the melody / sign of variations of the melody…
- Underlying harmony / chord sequences / root progressions
- Rhythmic placement / rhythmic cells and patterns
- Pauses / energy distribution and flow
- Drum sequences and loops, …
- Rhymes / phonetic flow / chant, …
- Instrumental timbres / arrangements and support / special effect schemes, …
- Macroscopic properties, such as mode, tonality, tempo[1]…
- etc…

[1] Such properties usually vary at a slow pace but they can acquire the status of structuring properties, if they create *patterns* at the working scale.

At this point, it is essential to underline a second key assumption behind the S&C approach : it is not the intrinsic *nature* of a musical property that confers to it the status of structuring property within a segment, but its *relative behavior* within that segment.

This assumption is not in contradiction with the fact that a particular musical dimension may frequently play a predominant role in conveying or signaling structural information for a particular music piece, or within a given genre. But considering that potentially <u>any</u> musical dimension can contribute to the inner structure of the musical segment makes it possible to approach a much wider range of musical contents within a common framework, in particular those for which the "usual" musical dimensions obviously do not prevail in their inner organization.

In that sense, we view the polymorphism of the S&C model as a very attractive feature. Conversely, the identification of the structuring properties constitutes a crucial step in the S&C analysis, as discussed further on.

4.3 Morphological elements

In this article, we will, in most cases, consider morphological elements (namely the x_{ij}) consisting of musical fragments of a few seconds (typically 2 bars). But this is not inherent to the approach and we will sometimes deal with shorter or longer units, without any loss of generality. Indeed, the principles of the S&C is not specific to any given time-scale and it is able to account for systems observable at several scales, as will be exemplified in some of our case studies.

Up to now, it has been implicitly assumed that the morphological elements are of equal size. Even though this will turn out to be often the case in practice, this is not an absolute constraint and situations occur where elements are of different sizes. In fact, the size can happen to be one of the properties which governs the system or which takes part to the contrast.

The ratio between the size of the segment and the (typical) size of the morphological elements composing the S&C will be called its (relative) *granularity*[1]. In the examples above, this value has actually been chosen equal to 4. This will remain the case in most of the musical cases studied in this article.

4.4 Syntagmatic relationships

As the carrier system is assumed to be easily self-deducible, decoding is made easier if the syntagmatic relationships *f* and g (and, to a certain extent, the contrast γ) are rather simple, in order to enable their direct detection from the musical content. This point of view is well developed in (Narmour, 2000), under the concept of *cognitive rules*.

Typical syntagmatic relationships can therefore be functions which operate on the notes :

- identity (exact or almost exact repetition) – denoted as id and $\widetilde{id}$ in the forthcoming text,
- chromatic transposition (constant shift in half-tones on the chromatic scale),
- diatonic transposition (constant shift in degrees on the current scale),
- mode or scale change (same degree(s) on a different scale),
- time-inversion (reverse order of notes),
- complementation (in various possible ways),
- etc…

but they may also be functions which operate on the amplitude, time or timbral dimensions such as :

- amplitude increase / decrease / silencing (i.e. zeroing the amplitude)
- fragmentation / augmentation / expansion (i.e. shortening, lengthening)
- extension / simplification (i.e. insertion / deletion of auxiliary musical material, such as ornaments)
- adjunction / suppression / change of instrument(s)
- etc…

The musical process is essentially multi-dimensional, but of course, within a structural segment, morphological elements do not exhibit close relationships with one another on all musical dimensions. Sometimes, they even appear to differ completely and when it is impossible to formulate a simple correspondence between two morphological elements, the syntagmatic function can be understood as an *in extenso* (re-)mapping from one element to the other, which we will denote *new*.

Finally, as an intermediate situation between *id* and *new*, we also consider the case where two morphological elements start alike (in their first half) but continue and end differently (in their second half). This *semi-id* function will be denoted *id'*.

The contrast function γ follows similar principles, except that it can be subject to much more variability than *f* and g, and still be identifiable as a contrast. Conversely, the special case when $\gamma = id$ (no contrast) leads to the plain realization of the carrier system.

Within a structural segment, one (or several) musical dimension(s) form(s) an obvious syntagmatic network of relationships and this can be viewed as creating the overall cohesion of the segment. It must also be noted (and it will be developed later) that additional systems usually develop at lower scales and granularities within the segment.

---

[1] Another term could have been *resolution*, but it would have been a source of ambiguity given the usual meaning of that word in music.

4.5 Morpho-syntagmatic patterns

In this section, we denote as a the primer (previously x_{00}), so as to simplify the notations.

Given the network of syntagmatic relationships f, g and γ, the unfolded sequence of morphological elements can now be written :

$$a \;\; f(a) \;\; g(a) \;\; \gamma(g(f(a)))$$

The use of the identity function id (i.e. repetitions) for f and/or g on some musical dimensions is rather frequent in music and leads to typical *patterns*, which we are inventorying and discussing in this section. We also consider situations where syntagmatic functions can be viewed as simple shift or symmetry operations, thus yielding a larger range of morpho-syntagmatic configurations.

• Repetition-based primary patterns

When one of the relations f or g within the system is either id or new, we can denote as b and c, distinct elements from a (and from one another). Such patterns turn out to be very easy to detect, as id (or $\widetilde{id}$) are undoubtedly the most obviously deducible functions within an S&C, whichever musical dimension is involved.

Generating all S&C corresponding to these situations yields the following 8 patterns (two of which being particular cases of the others, where $c = a$) :

f	g	Plain ($\gamma = id$)	Contrastive ($\gamma = new$)	
id	id	$aaaa$	$aaab$	
id	new	$aabb$	$aabc$	$aaba$
new	id	$abab$	$abac$	$abaa$

Table 1 : List of primary repetition-based patterns generated by the S&C model

The patterns listed in Table 1 indeed correspond to configurations frequently observed in music pieces, at different timescales ranging between a few seconds (sometimes even less) up to 25-30 seconds (or more). These patterns are usually considered to be strongly indicative of the structure of the passage, whether in classical pieces or in pop music.

They can be considered as unambiguous S&C patterns as it is unequivocally possible to determine the carrier system (column "Plain").

Patterns of Table I can easily be generalized to $\widetilde{id}$ (*almost-id*) and id' (*semi-id*) functions, straightforwardly yielding variants such as : $aa\tilde{a}b$, $aaa'b$, $ab\tilde{a}c$, $aba'c$, $a\tilde{a}bc$, $aa'bc$, etc…

• Repetition-based secondary patterns

Six other patterns can be obtained as the combination of 2 or 3 distinct elements (see Table 2). They can also be described as S&C, but they happen to be ambiguous : indeed, except for a few particular cases, it is not possible to tell whether the corresponding pattern is plain or contrastive. For instance, $abbc$, can be either contrastive or non-contrastive, depending on whether the function assumed to relate b to a also relates c to b in the same way. This may not be easy to arbitrate if f has a complex behavior.

As opposed to patterns from the primary set, secondary patterns listed in Table 2 are somehow weaker in the context of the S&C model, as, with a few exceptions, they tend to support less well-defined structural information.

f	g	Ambiguous (constrastive or plain)		
new	f	$abbc$	$abba$	$abbb$
new_1	new_2	$abcc$	$abca$	$abcb$

Table 2 : list of secondary repetition-based patterns generated by the S&C model
new_1 and new_2 denote two distinct new functions

• Shift-based patterns

In some cases, functions f and/or g can be viewed as a mapping by "translation" or "rotation" of one morphological element over the other (loudness increase, chromatic transposition, diatonic transposition, etc…), i.e. typically, a transformation that can be interpreted as a shift on a particular quantity (for instance, intensity, …) or scale (e.g. pitch, …). These correspond merely to what Narmour (2000) calls *iterative rules*.

In that case, the relation between the two elements can be denoted by an index such as a_1a_2, etc…, expressing their direct correspondence. In particular, when a sequence morphological elements forms a *regular progression*, it can be denoted as $a_1a_2a_3$, which corresponds to f being a shift and $g = f^2$. In the context of the S&C model, this can lead to a "plain" progression $a_1a_2a_3a_4$ or to a "broken" progression, such as $a_1a_2a_3b$.

Taking into account such shift configurations yields additional types of patterns, such as : a_1a_2bc, a_1ba_2c, or even ab_1b_2c and abc_1c_2 as generalizations of the repetition-based patterns of Tables 1 and 2. Shift functions may also be used to create a contrast (i.e. be used as the γ function).

• **Symmetry-based patterns**

Ultimately, it is worth mentioning transformations which can be understood as symmetry operations[1]. Symmetries may apply to melodic contours, chord progressions, intensity curve, and they relate to what Narmour (2000) calls *inversions* and *retrogrades.*

Denoting as a^* the corresponding operation, enables the description of specific patterns such as aa^*ab, aa^*bc, abb^*c, aba^*c, $abcc^*$, etc... Note that any symmetry is involutive (i.e. $a^{**} = a$) and symmetries can be (and often are) used as contrast functions γ.

• **Summary**

From this rapid study, it appears that the S&C model is able to encompass a wide variety of (more or less) familiar morpho-syntagmatic patterns under a unifying (and simple) framework; this being made possible by the 2×2 matrix scheme which governs the infrastructure of the model.

Moreover, it must be underlined that, in the general case, the S&C results from the synchronous realization of several such patterns over distinct musical dimensions, while, at the same time, some other musical dimensions may not follow any such patterns[2].

4.6 Correspondence with musicological concepts and theories

In this section, we discuss how the S&C model relates, and somehow combines, a number of musicological concepts : carrure, formal types, grouping, rule-mapping and prototypes.

• **Carrure**

There is a direct connection between the S&C model and the concept of *carrure* originally pointed out by Fetis (1830, pp. 60-62) as being used extensively by Mozart and many other composers after him. According to (Brennet, 1926) the *carrure*[3] in music is defined as *the symmetry established between portions of the musical phrase, so as to divide it into fragments of equal length. The term "square" is specifically used for melodic forms whose periods proceed by 4 or multiples of 4 : a 8-bar phrase split into two equal 4-bar parts ; a 16-bar phrase split into 4 4-bar fragments.*

Even though this definition is slightly ambiguous (and does not exactly correspond to our definition of "*square*"), it conveys the idea of an underlying "*squareness"* of the inner organization of prototypical musical phrases, which can be considered as a template. It suggests some sort of hypermetrical organization of musical segments based essentially on metrical proportions between their subdivisions.

In that sense, the S&C model, in its square form, can be viewed as a way to formulate the *carrure* of a musical segment, or more generally, its compliance with (or its deviation from) a "square" template.

• **Formal types**

When considering the structure of musical segments, Schönberg (1967, pp. 21-30), and after him Caplin (1998), define two types of inner organizations, referred to (by Caplin, pp. 9-12) as *formal types* : the *period* and the *sentence*. Both types are normatively 8-bar segments, even though they may last for 16 or even 32 bars. They begin with what Schönberg calls a "two-measure phrase" (Caplin, a "two-measure idea", or a "basic idea"), which occupies the first quarter of the segment (typically what we consider as the primer).

The difference between the period and the sentence lies in the way the repetition of the basic idea is handled.

- In sentences, the basic idea is repeated immediately so it is presented twice in a row, forming what Caplin calls the *presentation*. The second part of the sentence, the *continuation*, can either be the result of transformations (or formal processes)[4] of the presentation, or the presentation of new ideas (cf. function *new*).

[1] The term "conjugation" could also be used.
[2] For instance, in a structural segment, the harmony may go $abac$, the drums $aaab$ and the lyrics $abab$, while the melody goes $abcd$.
[3] Literally meaning nowadays the "build" of a person, in French.
[4] This resonates well with the concept of syntagmatic function introduced in this article.

- The period differs from the sentence in the postponement of the repetition. This is done using the introduction of what Caplin calls a “contrasting idea[1]” between the two occurrences of the basic idea, which normatively lasts a quarter of a period. The first half of the period is called the *antecedent*, and the second half the *consequent*.

A period may therefore be written as $abac$, with c being unspecified. As for the sentence, it may be written as $aabc$, with both b and c being unspecified.

Rather strikingly, the S&C model encompasses, under a single framework, these two major formal types of musical segment constructions commonly used to analyze the musical language in classical pieces. Beyond these 2 basic forms, the S&C accounts for any simultaneous combinations of them (on different musical dimensions) as well as other familiar types of sub-structures, in particular progressions (plain or broken).

Examples 1, 2 and 3, at the end of this article, illustrate instances of S&Cs in classical music pieces. The first two were selected because they are also analysed in (Caplin, 1998). Examples 4-9 illustrate other configurations which the S&C model is able to account for in recent pop compositions, showing the ability of the model to describe structural patterns across various musical genres, non-conventional musical dimensions, and at several scales at a time (Example 9).

• Grouping

As already mentioned in the introduction, the description of a musical passage in terms of S&Cs can be viewed as a grouping operation, based on similar conceptions as Lerdahl and Jackendoff (1983). Indeed, elements forming a S&C share privileged relationships, the existence of which is creating a sense of musical consistency of the whole segment under consideration, even if not all musical dimensions participate in the system.

However, whereas Lerdahl and Jackendoff’s conception relies on a tree-based hierarchy driven by affinities between strictly adjacent segments, the S&C model assumes a matrix scheme which is able to account for tight relationships between elements which may not be contiguous. Moreover, in the S&C approach, the grouping operation results from the possibility to identify jointly a set of properties forming a *system* (and a contrast), rather than being guided by hierarchical preference rules setting successive priorities for the fusion of adjacent segments (even though, in a number of cases, the result may be equivalent).

• Rule-mapping

The implication-realization model of Eugene Narmour (1990, 1992) for analyzing melody structures, later extended under the term of *cognitive rule-mapping* (2000) to a wider range of situations and to other musical dimensions is firmly grounded on similar premises as those of the S&C model : the essential role of rule-mapping processes in human cognition of music completely agrees with our conception of syntagmatic functions. Narmour also underlines that “*deduced rules generate strong expectations [… which] can be denied, delayed or overreached*”, pointing out the essential role of what we call the “contrast”.

Here, a significant contribution of the S&C model is to formalize the cognitive process in a matrix framework, thus embedding a wide variety of situations into a single generic scheme. This results in simpler formulations, wider generalization capabilities, and much more agility in handling (i) non-contiguous elements within a segment, (ii) multiple musical dimensions at a time and (iii) different timescales simultaneously.

As detailed in section 5, the S&C model does not only provide a generic and compact description of the cognitive relationships which develop inside a musical segment. It also induces an hypermetric description of the inner organization of the segment, which can be described in reference to a prototypical square shape, and which therefore yields an accurate structural metric signature.

• Prototype

As expressed in (Lamont, 2001), with respect to prototype theory in general, *“the prototype is viewed either as a particular privileged exemplar of a given category or as an abstracted central tendency, and similarity is a function of the distance between a given item and the prototype, measured in terms of common and distinctive features.*

This meets our conception of the S&C, as an abstraction which is essentially encapsulating, into a common formulation, a wide variety of well-known forms. It is seen as some sort of prototypic scheme, which does not explain all aspects of the musical content, but which can be used as a yardstick when describing the inner organization of structural segments. The segments may be more or less well fitted to the model, but the S&C can serve as a template to guide the identification of one or several plausible descriptions of their inner structure and to provide a means of comparing these descriptions, with respect to their goodness of fit.

Of course, a significant number of structural segments cannot be satisfactorily described by a purely square system and the next section introduces extensions of the S&C system to non-square configurations.

[1] In the present article, the “*contrasting idea*” of (Caplin 1998) corresponds to our concept of *opposition* (i.e. something “different”) but does not identify with that of *contrast* (something that departs from a logical system).

5. EXTENSIONS OF THE SQUARE SYSTEM

Up to this point, focus has been put on square systems, i.e. systems composed of 4 elements, which constitute some sort of canonical hypermetric configuration. This section presents how these principles can be extended to sequences of fewer (2, 3) or more (5, 6, 7, …) morphological elements, and how these sequences of various lengths derive from and/or generalize the prototypic square system.

We review in detail a set of typical configurations and we introduce a codification of the system organization, by specifying the *shape* of the carrier system, the possible deviations or irregularities which are observed in the realization of the system and the number of morphological elements impacted by the contrast.

This results in an inventory of approximately two dozen of typical shapes which can be used when describing the internal organization of a structural segment. In particular, this typology can be referred to for matching a musical segment with the S&C *shape* which best fits it, and which can therefore be viewed as some sort of structural meter signature.

5.1 Morpho-syntagmatic shape codification

In the forthcoming sub-sections, we will use the following notations, which will be further detailed in the text :

shape of the carrier system	$p \times q$			
insertion of a morphological element	$+1$		$\&1$	
suppression of a morphological element	-1		$\backslash 1$	
superposition / factorization of a morphological element (tiling)	$(-1+)$			
no contrast / contrast / double contrast / semi-double contrast	$:$	$.$	$..$	$\therefore$
shape at the immediately lower time-scale	$\frac{1}{2}\,(p' \times q')$			

Table 3 : Notations used to codify morpho-syntagmatic S&C shapes

We call *system size*, the number of morphological elements in the system, which is equal to the product of the 2 terms composing the shape (and more generally, to the result of the operation which codifies the shape).

In the case of square systems, the system shape is codified as in Table 4.1 and square system size is equal to 4 :

Configuration	General unfolded form	Shape codification
Square Carrier System (no contrast)	$x_{00}\ x_{01}\ x_{10}\ x_{11}$	$2 \times 2\ :$
Square System & Contrast	$x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}$	$2 \times 2\ .$

Table 4.1 : Unfolded form and shape codification of standard square systems

5.2 Dyadic sequences and bi-dyadic system

Even if we consider a timescale for which structural blocks are typically composed of 4 morphological elements, there may occur, at that timescale, segments consisting of 2 elements (i.e. of system size equal to 2), which we call dyadic sequences. They can appear as a repetition aa, a semi-repetition aa' or an opposition ab. They can be considered as a 1-dimensional system at that scale (1×2), but they can generally be described as a two-dimensional (square) system at the immediately lower scale, hence the notation : $\frac{1}{2}\,(2 \times 2)$.

The immediate repetition of a dyadic system naturally forms a non-contrastive square system : $aaaa$ or $abab$, which is frequent in some types of pop music. We call this a bi-dyadic system, which is a particular case of plain square system.

Configuration	General unfolded form	Shape codification
Dyadic repetition	$x_0\ x_0$	$1 \times 2\ :$ or $\frac{1}{2}\,(2 \times 2)\ :$
Dyadic opposition	$x_0\ x_1$	$1 \times 2\ .$ or $\frac{1}{2}\,(2 \times 2)\ .$
Bi-dyadic Square System	$x_0\ x_1\ x_0\ x_1$	$2 \times (1 \times 2)\ :$ or $2 \times \frac{1}{2}\,(2 \times 2)\ :$

Table 4.2 : Unfolded form and shape codification of dyadic and bi-dyadic systems

5.3 Triadic systems

Triadic systems (i.e. systems of size 3) can be viewed as a singular class of systems based on a single function *f*, used twice in a sequence, in the carrier form : $a\ f(a)\ f(f(a))$. The typical triadic carrier system can be written as : $a_1 a_2 a_3$

and its contrastive form is $a_1 a_2 b$ where $b \neq a_3$. In particular b can be quasi-identical to a_1 or a_2, yielding $a_1 a_2 a_1$, $a_1 a_2 a_2$ or $a_1 a_1 a_2$ as particular triadic patterns.

Configuration	General unfolded form	Shape codification
Triadic Carrier	$x_0\ x_1\ x_2$	1×3 :
Triadic S&C	$x_0\ x_1\ \bar{x}_2$	1×3 .

Table 4.3 : Unfolded form and shape codification of triadic systems

5.4 Truncated / elided square systems

Some triadic segments can also be analyzed as particular cases of *square* systems. Those correspond to situations where the segment can be understood as the realization of a square system for which one of the syntagmatic functions (f, g or γ) is the "*delete*" function.

This yields a sequence of morphological elements where one of the morphological elements can be denoted as nil (for instance : $x_{00}\ x_{01}\ x_{10}\ nil$), where nil denotes the total absence of musical substance, i.e. not even "silence", but "jump to what comes next in the piece", i.e. typically, the beginning of a new system, when the missing element is the contrast.

Configuration	General unfolded form	Shape codification
Truncated Square S&C	$x_{00}\ x_{01}\ x_{10}$	$(2 \times 2) - 1$.
Elided Square S&C (beginning)	$x_{01}\ x_{10}\ x_{11}$	$-1 + (2 \times 2)$.
Elided Square S&C (middle)	$x_{00}\ x_{01}\ \bar{x}_{11}$ $x_{00}\ x_{10}\ \bar{x}_{11}$	$(2 \times 2) \setminus 1$.

Table 4.4 : Unfolded form and shape codification of truncated or elided square systems

It appears to be particularly relevant to distinguish truncated or elided square S&C from triadic ones when there exist, somewhere else in the piece, the realization of the whole square system, to which the truncated system directly relates.

5.5 Pentadic sequences and systems

Pentadic configurations (i.e. sequences formed of 5 morphological elements) are, in practice, rather common in music. However, this is not in contradiction with the use of a square prototype model. Indeed, pentadic configurations can be quite straightforwardly related to a square system (which we call a *stem*) enriched by the insertion of an additional element (which we call an *affix*).

We distinguish the situations where the additional element is :

1. A *prefix*, i.e. the quasi-replication of the primer in first position, or the insertion of an element which forms a sub-system with the primer (denoted x_{00}^0 in its general form). Prefixes are observed in segments which begin with a false start, where some properties behave for instance like $aabac$, $aaabc$, etc…
2. A *suffix*, i.e. the quasi-replication of the contrast in last position, or the adjunction of an element which forms a sub-system with the contrast (denoted $\bar{x}_{11}^1$). Suffixes are very frequent and usually correspond to a repetition of the last element in the segment, yielding patterns such as $abacc$, $aabcc$, etc…
3. An *infix*, i.e. the insertion of an extra element within the S&C, usually (but not necessarily) in 4th position (i.e. just before the contrast). The infix itself can be either :
   a) the realization of the 4th element of the carrier system (x_{11}) before the actual contrast (thus postponing the contrast) – for instance, $ababc$, $aabbc$, …
   b) the quasi-replication of the previous element of the system (redundant infix) or the adjunction of an element which forms a sub-system with that previous element (denoted with a 1 superscript : x_{01}^1 or x_{10}^1) – typically $abaa^1c$, $aabb^1c$, …
   c) the presentation (in 4th position) of an element that relates to the forthcoming contrast, and which forms a sub-system with it (denoted with a 0 superscript as $\bar{x}_{11}^0$), which can be viewed as a pre-contrast – typical examples are $abac^0c$, $aabc^0c$, …
   d) the insertion of an "exotic" element (extraneous infix), creating some sort of parenthesis or diversion in the system's internal logic (denoted y), and yielding configurations such as $abayc$, $aabyc$, etc…

Table 4.5 inventories the various typologies of pentadic systems and Figure 8 illustrates S&C syntagmatic networks for a number of them. In fact, as can be seen on Figure 8, all these situations are covered by the introduction of a third morpho-syntagmatic function h which links the extra element to its most straightforwardly related neighbor (including $h = new$ in case 3.d, i.e. extraneous infix).

Configuration	**General unfolded form**	**Shape codification**
Square S&C with prefix	$x_{00}^{0}\ x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}$	$1 + (2 \times 2)$.
Square S&C with suffix	$x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}\ \bar{x}_{11}^{1}$	$(2 \times 2) + 1$.
Postponed contrast	$x_{00}\ x_{01}\ x_{10}\ x_{11}\ \bar{x}_{11}$	$(2 \times 2)\ :\ 1$.
Redundant infix	$x_{00}\ x_{01}\ x_{10}\ x_{10}^{1}\ \bar{x}_{11}$ $x_{00}\ x_{01}\ x_{01}^{1}\ x_{10}\ \bar{x}_{11}$	(2×2) & 1 .
Pre-contrast	$x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}^{0}\ \bar{x}_{11}$	
Extraneous infix	$x_{00}\ x_{01}\ x_{10}\ y\ \bar{x}_{11}$	

Table 4.5 : Inventory of unfolded forms and shape codifications of pentadic S&C

Pentadic S&Cs can thus be understood as a square S&C (called *stem*) extended by an "*insert*" operation, fully specified by the location of the infix and the relationship h which the infix exhibits with its most similar neighbor in the stem.

It may occur that the extra element does not show any strong relationship with either of its neighbors, in which case function h can be understood as the function new (see for instance example 10).

It may also happen that several stems are conceivable, or several interpretations of the infix are possible (for instance, to consider a redundant infix in 4th position also as a pre-contrast, because it exhibits relationships with *both* its neighbours – see example 11). Rather than a flaw, this should be considered as reinforcing the consistency of the sequence as a pentadic system, as several descriptions are able to account for the observed construction and therefore concur to the *same* grouping conclusion.

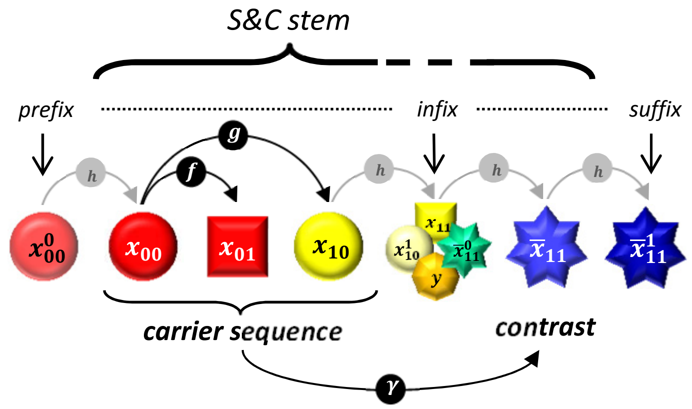


Figure 8 : Synthetic view of several pentadic configurations (unfolded forms) based on a square S&C stem

5.6 Hexadic sequences

In a number of cases, hexadic (6-element) sequences result from a set of 3 syntagmatic functions and their carrier system can be viewed as *rectangular* rather than square. Two categories of rectangular carrier systems must be distinguished depending on the relative span of the 3 syntagmatic functions. We call them *wide* versus *tall* and their carrier form is described in Table 5 below.

Configuration	**Unfolded form**	**Shape codification**
Wide hexadic (carrier form)	$a\ f_1(a)\ f_2(a)\ g(a)\ f_1\big(g(a)\big)\ f_2\big(g(a)\big)$	2×3 :
Tall hexadic (carrier form)	$a\ f(a)\ g_1(a)\ f\big(g_1(a)\big)\ g_2(a)\ f\big(g_2(a)\big)$	3×2 :

Table 5 : Wide and tall hexadic carrier system

However, a number of hexadic sequences are better explained as square S&Cs with multiple affixes. They can be viewed as further extensions of pentadic systems, with 2 independent affixes, for instance, $1 + (2 \times 2)$ & 1. But they may also be the result of a double infix nested in the square system, itself forming an independent dyadic sequence : (2×2) & (1×2).

Configuration	**Unfolded form**	**Shape codification**

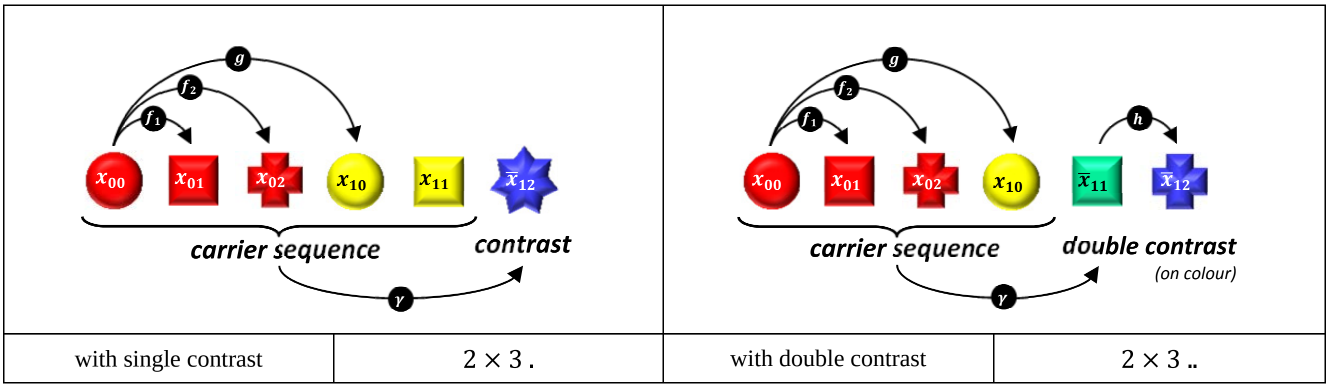
Wide hexadic S&C	$x_{00}\ x_{01}\ x_{02}\ x_{10}\ x_{11}\ \bar{x}_{12}$	2 × 3 .
Wide hexadic S&C with double contrast	$x_{00}\ x_{01}\ x_{02}\ x_{10}\ \bar{x}_{11}\ \bar{x}_{12}$	2 × 3 ..
Tall hexadic S&C	$x_{00}\ x_{01}\ x_{10}\ x_{11}\ x_{20}\ \bar{x}_{21}$	3 × 2 .
Tall hexadic S&C with double contrast (ambiguous with plain square + dyad)	$x_{00}\ x_{01}\ x_{10}\ x_{11}\ \bar{x}_{20}\ \bar{x}_{21}$	3 × 2 ..
Tall hexadic S&C (blues-like form) with semi-double contrast	$x_{00}\ x_{01}\ x_{10}\ x_{11}\ \bar{x}_{20}\ x_{21}$	3 × 2 ∴
Plain square system followed by dyadic sequence	$x_{00}\ x_{01}\ x_{10}\ x_{11}\ y_0\ y_1$	$(2 \times 2) : \frac{1}{2}(2 \times 2)$.
Square S&C with nested dyadic sequence	$x_{00}\ x_{01}\ x_{10}\ y_0\ y_1\ \bar{x}_{11}$ $x_{00}\ x_{01}\ y_0\ y_1\ x_{10}\ \bar{x}_{11}$	$(2 \times 2)\ \&\ \frac{1}{2}(2 \times 2)$.

Table 4.6 : Unfolded form and shape codification of several hexadic S&C

Note that a plain square system followed by a dyadic sequence is inherently ambiguous with a tall hexadic with double contrast (in fact, both of them form a triadic sequence $a_1 a_2 b$ at the immediately upper scale). By convention, we will preferably opt for the second solution, except if there is a strong and obvious relationship between 5 of the 6 elements taken as a whole. In particular, this will be the case with blues-like musical segments where the 6^{th} element usually relates strongly to elements 2 and 4, while the contrast impacts mostly the element in 5^{th} position.

Figures 9a and 9b illustrate the various hexadic sequences listed in Table 5.6. Example 12 provides an instance of a plain wide hexadic S&C (2 × 3 :) and example 15 contains, towards the end, a plain square system followed by a dyadic sequence (segments labeled E and F).

Wide hexadic S&Cs



Tall hexadic S&Cs

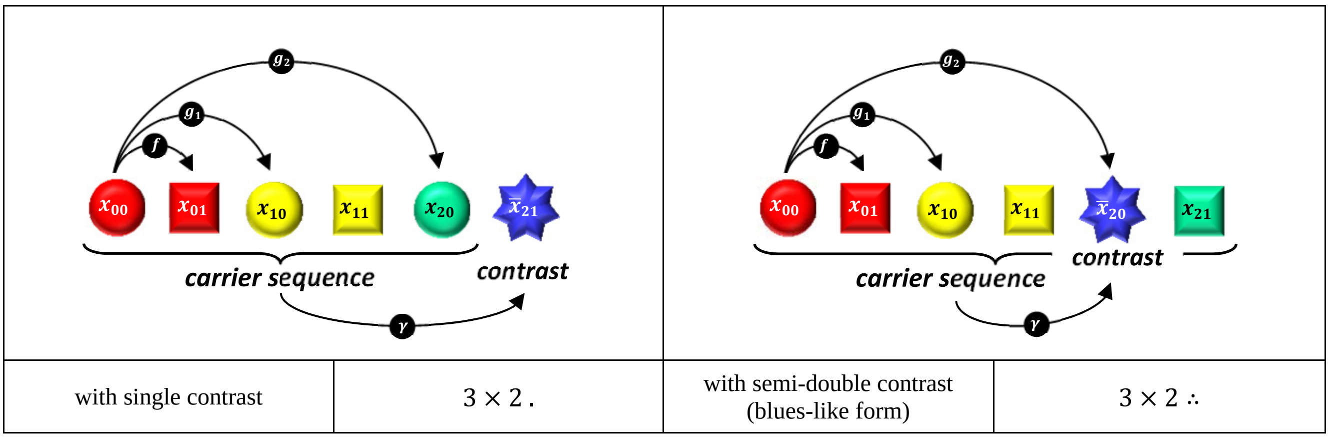
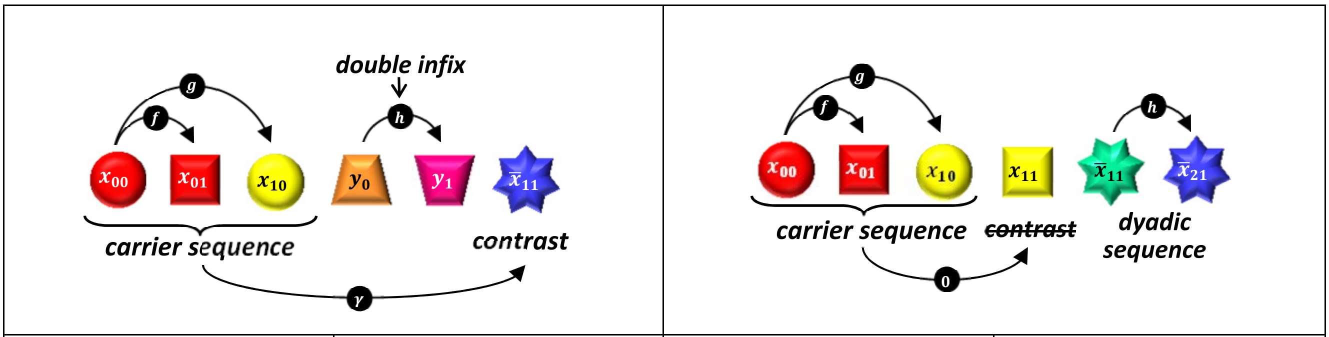


Figure 9a : Illustration of hexadic systems together with their denomination and shape

Hexadic sequences deriving from square systems



Square S&C with nested dyadic sequence	$(2 \times 2)\ \&\ \frac{1}{2}(2 \times 2)\,.$	Plain square system followed by a dyadic sequence	$(2 \times 2) : \frac{1}{2}(2 \times 2)\,.$
		This configuration is ambiguous with a tall hexadic with double contrast	

Figure 9b : Illustration of hexadic sequences together with their shape codification

5.7 Larger systems

In line with the logic now established in the previous sub-sections, heptadic (7-element) segments can be analyzed as various affixed versions of square or rectangular systems. We leave it to the reader to envision constructions such as : $1 + (2 \times 2)\ \&\ \frac{1}{2}(2 \times 2)$, $(2 \times 3) + 1$, $(3 \times 2)\ \&\ 1$, $(2 \times 2)\ \&\ (1 \times 3)$, $(2 \times 2)\ \&\ (2 \times 2 - 1)$, etc…

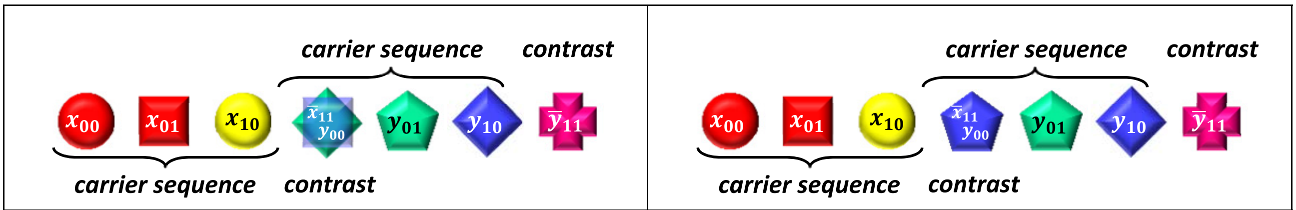
Similarly, some 8-element segments can be described as complex irregular forms deriving from smaller carrier systems. But some may also result from *cubic* systems $(2 \times 2 \times 2)$, i.e. systems based on properties evolving along three distinct musical dimensions and/or variation cycles. In general, however, cubic systems can be approximated as two successive square systems, by neglecting the syntagmatic function with the longest span. Alternatively, they can also be approached as a single square system at the immediately upper time-scale, by grouping neighboring elements two by two.

Nonadic systems (3×3) are perfectly conceivable, but they turn out to be quite rare, at least at the 2-bar morphological granularity.

5.8 Tiling

This section would not be exhaustive without briefly mentioning an occasional yet important configuration resulting from cases where 2 successive systems overlap, thus creating a situation where the contrast of a given S&C and the primer of the next S&C are either superposed (played/heard at the same time) or merged into a single element (thus functionally acting both as a contrast and a primer). The two typical forms of tiling are depicted on Figure 10.

Tiling configurations



Surface tiling	$2 \times 2\,.(-1+)\,2 \times 2\,.$	Functional tiling

Figure 10 : Two types of S&C tiling (illustration)

These two configurations, which we call respectively *surface tiling* and *functional tiling*, can be understood as the use of a common time-interval and/or a single morphological element shared between the 2 systems. This configuration sometimes appears as the most efficient way to describe a sequence of morphological elements, rather than introducing systems with more irregular shapes. Situations of (partial) tiling can be found in Example 15.

5.9 Simplicity principle

S&C configurations such as those inventoried in this section boil down to a limited set of shapes which characterize and reflect some high-scale hypermetrics governing the inner organization of musical segments, as noted in (Bimbot, Deruty, Sargent, Vincent, 2011). Similarly to meter signatures, they can be binary, ternary, quaternary, … regular or irregular, simple or compound, and realized in a complete or incomplete form.

The S&C model thus makes it possible to describe and codify the inner organization of structural segments. However, when facing a musical passage, several options may occasionally arise which are equally acceptable in terms of S&C description.

As a general principle, we assume that, at a given time-scale, priority should be put on systems whose size ranges between 2 and 6 morphological units (with a strong priority on square stems), in order to favor comparability across segments. Therefore, by convention, smaller or larger systems will preferably be merged or split, to concentrate the distribution of segment sizes around the target time-scale (while keeping the option to do otherwise, if there are very good reasons to do so).

Similarly if several options leading to different segment boundaries seem equally valid, we recommend opting for the simplest one.

Within segments, it may happen that the hypermetric shape may differ from one musical dimension to the other (for instance $(2 \times 2) + 1$ *vs* (2×2) & 1, as in example 11). But an essential constraint overrides the ambiguity of these situations : the borders of the systems are *synchronized*, which ultimately reinforces the consistency of the segment as a single unit.

In quite a number of cases (particularly in pop music), most segments at a given scale have a comparable system size across the piece, and the piece can be viewed as being built upon a well-identifiable "structural pulsation period" (Bimbot, Le Blouch, Sargent, Vincent, 2010). In other pieces, system shapes and sizes vary more freely along the piece, creating the perception of an irregular "structural rhythm".

6. THE S&C DESCRIPTION AS A MODEL MATCHING TASK

In this section, we develop the criteria and methods which can be used for decomposing a structural segment or a full passage in terms of S&C, and we thus formulate principles on which our analysis and discovery process is based.

We discuss how the description of a musical segment in terms of S&C can be viewed as a coding scheme and therefore leads to a formulation of the S&C analysis process as a model matching problem aiming at maximizing a compression gain. In this framework, we highlight the essential role played by the primer.

We show how the model matching viewpoint is useful to guide the arbitration between several S&C hypotheses which may happen to be in competition for modeling a given structural segment, but also for determining the decomposition of a long passage or an entire piece into successive S&Cs.

In particular, we underline that :

- In a structural segment, the various musical dimensions may not follow the *same* structural pattern (some may for instance behave as $abac$, others as $aabc$) but they all contribute to the S&C, provided they exhibit a synchronous behavior.
- In general, *not all musical dimensions* do participate to the S&C, but only a subset of them, while the other musical dimensions exhibit an unstructured behavior. This requires a step of selection of variables.
- Segmentation of a passage using the S&C model results from a joint optimization of successive S&C over the whole passage.
- S&C patterns generally co-exist *at several timescales* simultaneously in a given structural segment and this contributes to the reinforcement of the consistency of the segment.

6.1 Law of parsimony

Referring to simple examples such as those introduced in section 2, it appears that the determination of the underlying S&C within a set of elements is based on the inference of :

a) the necessary and sufficient set of properties which are required to describe the elements,
b) the simplest set of functions that is needed to formulate the relations between the elements,
c) the precise subset of properties which are relevant to characterize the system and its contrast.

Humans tend to achieve intuitively the joint estimation of these properties and their relationships by eliciting, between several acceptable descriptions, the one which seems the most obvious and straightforward[1]. This cognitive strategy relates to the Ockham's razor principle (also called the *law of parsimony*) (Ockham, 1323), which, loosely speaking, assumes that, among several possible ways to describe a same set of observations, the preference tends towards that which is the simplest (in terms of prior assumptions and complexity).

[1] Clearly, there can be ambiguous cases for which several distinct descriptions are comparably obvious and straightforward.

In the field of engineering sciences and Information Theory, the Minimum Description Length (MDL) approach (Rissanen, 1978) is a particular instance of this principle, based on the fundamental idea that "*any regularity in a given set of data can be used to compress the data, i.e. to describe it using fewer symbols than needed to describe the data literally*" (Grünwald, 1998).

Asymptotically, the upper compression bound of a set of data corresponds to the Kolmogorov complexity of the data (Kolmogorov, 1963), i.e. the shortest *program* that outputs the data. MDL restricts the search of the optimal compression scheme to a subset of allowed *codes*, called the model class, which is chosen to be reasonably efficient, whatever the data at hand.

The MDL scheme directly relates to probability theory through the correspondence between codes and probability distributions and can be viewed, under certain hypotheses, as asymptotically equivalent to Bayesian inference (MacKay, 2003).

6.2 S&C model matching criterion

Based on these principles, the adequacy of an S&C in describing a musical passage can be formulated as an information-theoretic criterion which expresses how much compression gain can be achieved by modeling the passage as an S&C rather than describing it literally.

In this context, the musical passage $S = x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}$ plays the role of the literal data, the S&C model is the model class (the "coding scheme") and the S&C description of the passage, $M = (x_{00}, f, g, \gamma)$ is the compressed data.

Let's assume that the quantity of information (i.e. the number of bits) which is required to describe S literally writes :

$$q_0(S) = q_0(x_{00}) + q_0(x_{01}) + q_0(x_{10}) + q_0(\bar{x}_{11})$$

where $q_0(x_{ij})$ therefore denotes the quantity of information required to describe each element of the system. We will consider that each of these terms has a typical order of magnitude of Q.

When compressed as M, the quantity of information needed to describe S becomes :

$$q_M(S) = q_0(x_{00}) + q_M(f) + q_M(g) + q_M(\gamma)$$

where $q_M(\theta)$ denotes the quantity of information required to encode a given function θ.

The compression gain achieved by the S&C model can therefore be expressed as the difference between these two description costs, namely :

$$\Delta q(S|M) = q_0(S) - q_M(S)$$

For a given sequence S, the most favorable description M^* will therefore be chosen as the one which minimizes $q_M(S)$ (i.e. maximizes $\Delta q(S|M)$).

It is not within the scope of this article to develop explicit expressions of q_0 and q_M. In fact, in human-based analysis, the most likely description M^* is estimated intuitively by the analyst. However we do show that, if several hypotheses are in competition, they can be empirically gauged and ranked, taking into account a number of "common sense" considerations, which ensue from the formal expression of Δq :

- When one of the functions (f, g or γ) is of the class "*new*", the associated term in q_M should be considered as equal to the corresponding one in q_0 (i.e., typically Q) : indeed, it is reasonable to consider as equivalently complex and costly either (i) to update completely the new element in the S&C model or (ii) to describe it literally.
- Conversely, as soon as there exists (at least) one musical dimension for which a simple S&C description of S is admissible, this entails some compression gain in $\Delta q(S/M)$. In particular, as soon as one of the two functions f or g are cognitively simple, i.e. for instance, for any sentence-like ($aa'bc$) or period-like ($aba'c$) behavior of a musical dimension within a segment, the low value (say ε) of the second or third term in $q_M(S)$ is sufficient to create a significant gain in Δq.
- The existence of synchronized compact descriptions of the system on several musical dimensions simultaneously all contribute to increase the compression gain (and therefore to reinforce the consistency of the S&C), even if the systems show different patterns on the considered dimensions (for instance, $aba'c$ on some musical dimensions and $aa'bc$ on other ones). This point is further detailed in the next section.
- Whatever type of function is considered to evaluate $q_M(S)$, it can be assumed that the values taken by q_M should be ranging as depicted in Figure 11.

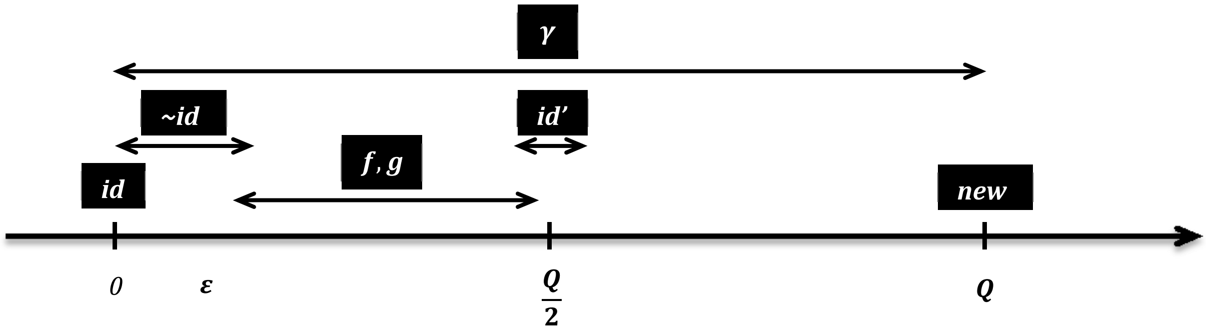


Figure 11 : Relative range of values taken by q_M for various syntagmatic functions

In particular, it seems reasonable to assume that admissible syntagmatic functions for modeling a morphological element as the transformation of another one should result in a significant compression ratio (at least a factor of 2).

- Note that the *primer* element x_{00} plays a key role in the compression criterion, but this role is indirect : its intrinsic complexity does not impact Δq, but the possibility to describe more or less compactly the properties of the *forthcoming elements* as simple and straightforward functions of that primer is determinant for the overall value of Δq.

6.3 Selection of variables

The S&C description process indirectly relies on the selection of the musical dimensions over which it is considered to be resting. Indeed, the model-matching approach can be viewed as operating, jointly to the estimation of the syntagmatic functions, a *selection of variables* over which the most compact description can be achieved.

Let's denote as $\mathbb{S}$, the subset of structuring properties of segment S, i.e. the properties which are considered as forming the carrier system of S and as $\overline{\mathbb{S}}$ all the other (non-structuring) properties. Criterion q_M therefore decomposes into two terms, corresponding to the quantity of information required to encode S by the S&C description in each separate subspace :

$$q_M(S) = q_M(S|_{\mathbb{S}}) + q_M(S|_{\overline{\mathbb{S}}})$$

Clearly, while the first term in the sum, $q_M(S|_{\mathbb{S}})$, is smaller than its literal counterpart $q_0(S|_{\mathbb{S}})$ because of the existence of simple syntagmatic relationships along the dimensions of $\mathbb{S}$, the second term, $q_M(S|_{\overline{\mathbb{S}}})$, can be assumed as being of the same magnitude as $q_0(S|_{\overline{\mathbb{S}}})$, as no compression can be expected from variables which vary unrelatedly with the model.

As a consequence, the compression gain actually writes as :

$$\Delta q(S|M) = q_0(S) - q_M(S|_{\mathbb{S}})$$

and its decrease is entirely attributable to the variables living in subspace $\mathbb{S}$. Conversely, matching a segment to an S&C model provides simultaneously an estimation of $\mathbb{S}^*$, i.e. the (largest) subset of variables responsible for the compression gain (conditionally to description M).

In general, not all musical dimensions may form S&C patterns in a structural segment, but only a subset of them, while the other ones vary in a non-organized way. For instance, the lead part played by the central instrument during a solo within a pop (or jazz) piece may temporarily break free from any structured patterns and the organization of the segment may be only governed (weakly) by the underlying harmonic progression. Similarly, in many rap pieces, the vocal lead is obviously unstructured, neither in terms of melody, nor of rhythmic pattern, and so on…

Nevertheless, there is usually at least one musical dimension (which may vary from one structural segment to the next one) for which some S&C pattern can be identified, thus indicating a quite clear hypermetric inner organization within the segment.

6.4 Segmentation into S&Cs

The essential role of the primer element x_{00} in the determination of an S&C becomes particularly obvious when dealing with the task of segmentation, i.e. grouping morphological elements into successive segments within a long passage (including the case of an entire music piece), and thus determining a structural segmentation where each and every segment's inner organization matches as much as possible the S&C model.

We restate here that we assume to be working around a given timescale, i.e. that the morphological elements consist of a pre-determined number of bars (typically 2 bars). It is also assumed that a preference is put on square systems but that some of them may be triadic, pentadic, hexadic, etc…

This extra degree of freedom creates a significant factor of complexity, because the *boundaries* of the successive S&C need to be optimized *jointly* with the S&C themselves. However, this turns out to be tractable because valid boundaries result from a simultaneous identification of the contrast of the finishing segment and of the primer of the forthcoming one.

As an illustration, let's consider the sequence of figure 12, in which, we want to consider (and arbitrate between) two segmentation hypotheses : $H_1: [x_1x_2x_3x_4][x_5x_6x_7x_8x_9] = S_1^1S_1^2$ versus $H_2: [x_1x_2x_3x_4x_5][x_6x_7x_8x_9] = S_2^1S_2^2$. In other words, the question is whether a segment border should be put *before* x_5 or *after* x_5.

x_1 x_2 x_3 x_4 x_5 x_6 x_7 x_8 x_9

Figure 12 : An unsegmented sequence of 9 morphological elements

Under hypothesis H_1, sequence $S_1^1 = [x_1x_2x_3x_4]$ forms a plausible square S&C, with a "shape" contrast in 4th position. Under hypothesis H_2, sequence $S_2^1 = [x_1x_2x_3x_4x_5]$ constitutes a plausible pentadic S&C, with x_4 acting as an infix and x_5 as the contrast. However, under H_2, $S_2^2 = [x_6x_7x_8x_9]$ appears as a clear square S&C with a "color" contrast, but it seems very difficult to find an economical decomposition of $S_1^2 = [x_5x_6x_7x_8x_9]$ as a pentadic system. Indeed, whereas x_4 can be a considered as an acceptable contrast in S_1^1, x_5 does not appear as a reasonable primer (nor prefix) in S_1^2. Therefore the preference goes for hypothesis H_2.

This arbitration process can be understood as the empirical minimization of the global description cost $q_M(S_i^1) + q_M(S_i^2)$, over the two hypotheses ($i \in \{1,2\}$). In this example, $q_M(S_1^1)$ and $q_M(S_2^2)$ are of comparable magnitude but $q_M(S_1^2) \gg q_M(S_2^1)$.

The preference for a segment boundary arises from the identification of a plausible contrast *before* the boundary but first and foremost, of a reasonable primer just *after* the boundary : the primer appears simultaneously as an element which is *poorly suited to the previous S&C* (typically, it is located after the contrast and shows no clear syntagmatic relationship with it) <u>and</u> an element which is *essential to describe* the forthcoming items in *the next system* on the basis of compact morpho-syntagmatic relationships

As was noted in (Bimbot, Le Blouch, Sargent & Vincent, 2010) structural segments are perceived as both *autonomous* and *suppressible* within the musical flow because their elements form a consistent set of their own (property of autonomy) and none of these elements are necessary to explain the elements of neighboring systems (property of suppressibility).

The joint estimation of successive S&Cs greatly facilitates the disambiguation of structural boundaries and generally conditions the global segmentation towards quasi-uniqueness (even when the internal structure of the segments is locally complex or ambiguous). Examples 15 and 16 illustrate the segmentation of two passages by means of S&C analysis and show how the method can locate and characterize accurately the successive structural segments at or around a given timescale.

6.5 S&Cs at differents timescales

S&Cs detectable around a given timescale are usually exhibiting simultaneously synchronous *sub-systems* at lower timescales, and they may themselves be embedded in larger *super-systems*. When working at a given timescale, the existence of these phenomena are useful to comfort the consistency of a segmentation hypothesis (for example, the succession of two sentences starting by the same presentation, but two distinct continuations, constitutes a period at the immediately upper scale). These multi-scale hierarchies reveal an important additional facet of music structure and modeling the interactions between S&Cs at different timescales is a topic that deserves to be studied and described in more detail in a future publication.

Altogether, a consistent structural segment in the framework of the S&C model fits well with the general definition given by Spencer & Temko (1988, p. 31) : "*a major structural unit perceived as the result of the coincidence of relatively large numbers of structural phenomena*" which, as we have highlighted it in this section, usually occur in several musical dimensions and at different time-scale simultaneously. In our conception, this corresponds to a tight network of multi-dimensional and multi-scale syntagmatic relationships.

7. COMMUNICATION FUNCTIONS OF THE CONTRAST

Communication is a very broad area which covers a variety of aspects dealing with the exchange of information, messages or thoughts, and there is no doubt, from this definition, that music is a communication process. In this section, we introduce and discuss what we view as various communication *functions* played by the contrast within structural segments.

We first explain that, within the framework of digital communication sciences, the contrast element can be interpreted as a logical (or digital) *modulation* of the last element of the *carrier* system. We then argue on the role of the contrast as constituting some sort of musical *punctuation mark* within the musical flow. Finally, we discuss the interpretability of the contrast as a *surprise* element that creates a *punch-line* in the musical narration.

The relationship between *information*, expectation and surprise in music is addressed in several work, including (but not limited to) Meyer (1967), Narmour (1977), Abdallah & Plumbley (2009). In this section, we focus more specifically on the role of the contrast as regards its *communication* functions, being conscious that both aspects are intimately linked.

Following the classical communication scheme, we assume, when needed, a source (the composer), a receiver (a listener), a message (the musical piece or passage) represented by a particular code (a system of information units) and transmitted on a channel (any information medium which conveys the musical content of the piece/passage).

Under this viewpoint, the S&C description of a segment can be viewed as a code-word describing the organization and the content of the musical message (under the assumption that the S&C model acts as a class of codes). In that context, the contrast plays a double role : that of being part of the musical message and that of acting as a delimiter within the code.

7.1 The contrast as a digital modulation

In the field of electronic engineering and telecommunications, the concept of *modulation* is commonly defined as the process of varying one or more properties of a known *carrier signal* (usually a periodic waveform) with an unknown *modulating signal* conveying information which fluctuates at a lower frequency than that of the carrier.

In analog communications, the modulation is applied continuously to the carrier signal as a function of the information content. In digital communications, the communication signal is usually a discrete process, and modulation consists in changing at specific instant the values of the bits constituting the code-words. In both cases, retrieving the modulating information from the composite signal is called *demodulation*.

As pointed out briefly in sections 2 and 3, the sequence $S_0 = x_{00}\ x_{01}\ x_{10}\ x_{11}$ can be viewed as a carrier signal, this signal being composed of a small number of distinct states for each musical dimension. The set of states and the state sequence are constructed at the source level by alternating in an organized (and supposedly decodable) way the properties of the morphological elements $x_{00}\ x_{01}\ x_{10}$ along the different musical dimensions.

The fact that the last element of the carrier sequence S_0 is fully predictable (i.e. redundant) provides the opportunity to insert additional information by modifying one or several properties (bits of information) of x_{11}. This yields the contrasted sequence $S = x_{00}\ x_{01}\ x_{10}\ \bar{x}_{11}$, where $\bar{x}_{11}$ (and therefore function γ) can be understood as a *digital modulation*[1] of the last element of S_0.

From the viewpoint of the receiver, the carrier is reconstructed by inferring the functions f and g from the sequence $x_{00}\ x_{01}\ x_{10}$ (this task is easier if f and g are cognitively simple). When the last element of the system occurs ($\bar{x}_{11}$), it is therefore possible for the receiver to analyze its deviation from the expected carrier element (x_{11}) and therefore to *demodulate* the contrastive properties (denoted $\bar{x}_{11} \div x_{11}$ in section 3.2) and to recover the extra information conveyed by the contrast.

In this scheme, as opposed to a conventional modulation-demodulation scenario, the receiver does not know the carrier signal in advance and must reconstruct it, for each successive segment, by deduction based on observed *regularities* and *redundancies* of the musical properties in the various musical dimensions. Note also that the contrast is just one particular form of modulation of the carrier system which does not preclude other non-systematic and slowly-varying evolutions of its properties (such as the texture in S&C #2 of section 2.4).

It is worth noting that, for a (square) S&C of length *n* bars, properties of the 2×2 square carrier tend to alternate at a period either equal to $n/4$ or to $n/2$ whereas the contrast occurs (only) at time intervals of n, i.e. at a longer period (lower frequency) than the carrier. This is entirely consistent with a modulation-demodulation scheme.

Of course, assuming that the internal organization of structural segments in music is governed by a carrier-modulation process should be understood as an abstraction which does not reflect all the aspects at work in the process of musical composition or listening. Nevertheless, we believe that relating the S&C model to digital communication principles provides an interesting explanation of some structural aspects of conventional music, which may provide openings towards new techniques in music analysis, processing, content description and information retrieval.

7.2 The contrast as a punctuation mark

The elements forming human communication messages as well as digital data stream are often grouped into larger units which constitute blocks of information. Either these blocks can be assumed to be of a fixed-length, or they must be delimited (explicitly or implicitly) by some easily identifiable boundary markers. This is typically the role of punctuation marks in natural human language, which serve to structure logically the linguistic message.

Whereas, in written communication, punctuation marks are inserted explicitly in the textual content as specific signs, they take a very different form in oral communication : in spoken language, the logical organization of the discourse is mainly rendered by *prosodic* markers at the end of logical groups of words : introduction of pauses, significant modifica-

[1] It is essential to underline that this particular use of the word *modulation* departs from its usual meaning(s) in the field of music.

tions of the intensity, the syllable flow or the vowel durations, inversion of the shape of the intonation contour, etc… Note that these prosodic modifications are realized simultaneously with the articulation of phonemes, i.e. they are embedded within the other levels of the linguistic message.

A rather striking parallel can be established between these prosodic processes and those occurring in the musical flow within the S&C. Indeed, similarly to a prosodic strategy, the contrast manifests itself as the last item of a sequence by a significant deviation of its musical properties from their current course, i.e. from reference properties that have been established beforehand in the preceding parts of the segment. Dually, it can be considered that spoken language uses mechanisms which consists in modulating its "musical" dimensions in order to mark the logical organization of the discourse, in the same way as the contrast element does in a S&C system.

In addition, note that, as for spoken language :

- the S&C segment boundary falls *after* the contrast,
- the contrast does not only mark the end of the current S&C but also announces the advent of a new one,
- the contrast is embedded in the musical discourse in a similar way as is the case for prosodic markers.

From these strong similarities, we feel that it is arguable to consider the contrast in an S&C as acting as a punctuation mark within the musical flow.

7.3 The contrast as a narrative surprise

A quick search for the meaning of the word *surprise* in various dictionaries returns definitions such as "an emotional state experienced as the result of an unexpected event", "the difference between expectations and reality", "the gap between our assumptions and expectations and the way that those events actually turn out" or "the end result of predictions that fail"[1],[2].

Creating surprise in a narration or, more generally, in any interactive situation, is indeed some sort of *art*. It plays an important role in novels, in plays, in jokes, in shows, in social events,… and of course, in music. Needless to insist on how well the above definitions of *surprise* do indeed apply to the concept of contrast as developed in this article.

More precisely, the contrast is in fact a particular type of surprise. Generally it does not constitute a completely novel and unforeseeable event in the musical flow. Firstly, the location of the contrast is generally rather predictable, as it can only occur after the presentation of the carrier system. Secondly, the contrast tends to share similarities on some musical dimensions with the carrier system and therefore is not totally new in every respect.

Therefore, the contrast must be understood as a phase of the musical narration where an unknown quantity of surprise is expected to happen, because a sufficient amount of information has been disclosed (in the carrier system) for the contrast to take its full *value*. In this respect, it is worth quoting Carl Philip Emmanuel Bach (Bach, 1753): "*Embellishments are best applied to those places where the melody is taking shape, as it were, or where its partial, if not complete, meaning or sense has been revealed.*"

Depending on the familiarity of the listener with the musical genre of the piece, the contrast may sound either quite surprising or very conventional. The nature and the intensity of the contrast also influence its perception by the listener. Ultimately, the surprise effect can come from the non-realization of a contrast (i.e. the realization of a plain carrier system), the surprise then resulting from… the absence of surprise.

It is not within the scope of this article to classify or qualify the narrative effects of a contrast with respect to its nature. Contrasts are bound to be appreciated very variably across listeners, according to their own aesthetical expectations. Indeed, listeners may judge a particular S&C as either sophisticated or common-place, subtle or disappointing, funny or boring, extraordinary or inept, surprisingly surprising or pathetically predictable, etc… Moreover, many of the surprising effects in music also come from other types of strategies.

But here we pinpoint the fact that the contrastive part of the S&C creates a recurring state of expectation which can be interpreted as analog to the punch-line of a story, or of an episode of a story, where, after having involved the audience into a given situation (the carrier system) the story-teller releases additional information which concludes (and resolves) the narration by creating a surprise built on material which is compatible with the beginning of the narration but which deviates from the outcome that would be the most in line with the context.

Example 13 provides an example where, in the context of a well-known comedy, the inherent contrast of a musical segment is reinforced by additional acting means in order to strengthen the punch-line effect of the last element. Conversely, example 14 exemplifies a musical excerpt which is not analyzable in terms of S&Cs and which clearly (and, as one may assume, deliberately) appears as a continuous flow of unstructured musical information.

---

[1] Note that surprise can be pleasant, unpleasant or simply neutral.

[2] Surprise must be distinguished from *novelty*, i.e. "the quality of being new, original or unusual".

7.4 Summary

In this section, we have identified various communication functions which can arguably be related with the contrast within a S&C.

As an element that punctuates the flow of a musical segment, the contrast of the S&C model can therefore be viewed as some sort of *morpho-syntagmatic cadence*[1], i.e, a way to signal weak (temporary) or strong (definite) closure within the musical narration by resolving a situation of expectedness created by the presentation of the carrier elements.

Future research may (or may not) comfort these considerations and better define their scope, their limits, and their conjunction with other levels where musical structure manifests itself.

8. POTENTIAL APPLICATIONS AND IMPLEMENTATIONS OF THE S&C MODEL

Even though the focus of this article is to present the S&C model in complete disconnection with any particular applicative context, it is worth mentioning briefly that the model appears to have a wide range of potential applications in computer sciences, in applied musicology and at the intersection between both.

In the field of music analysis, the S&C model has initially been exploited to characterize not only the internal organization of structural segments but also to guide the structural analysis of entire music pieces. In this context, we have recently introduced a methodology for decomposing and labeling a given music piece into structural segments (Bimbot, Deruty, Sargent & Vincent, 2012). The S&C model is called for to determine an unequivocal structural segmentation (at a given time-scale) but it is also used to assign "semiotic" labels to segments which characterize their similarities : two segments built from the same carrier system are associated to the same label, and differences pertaining to the contrast or to the surface realization of the system are denoted in different ways.

An interesting perspective that arises from the automatic analysis of large music datasets is that of building statistical models of music accounting for musical structure at different scales. Such models may be used in the field of MIR and computer-assisted musicology for, e.g., automatic music style classification or automatic detection of music pieces with similar structure. They may also be used to automatically generate shortened, extended or remixed versions of a given piece by reordering its structural segments in a new way. They may be considered as helpful to improve statistical music composition algorithms based on statistical chains (such as Markov chains) which so far mostly account for the short-term dependencies of music (Miranda, 2001) : the S&C paradigm could enable these models to generate entire music pieces with a more realistic structure.

Models of music structure are indeed useful for a range of MIR and music signal processing problems. Knowledge of the structural segments composing a music piece has been used to increase the accuracy of chord and tempo estimation (Mauch, Noland & Dixon, 2009 ; Dannenberg, 2005) or to estimate the short-term power spectrum of the accompaniment for singing voice separation (Liutkus, Rafii, Badeau, Pardo & Richard 2012), but the internal organization of these segments remains to be exploited. Other MIR problems such as polyphonic pitch and drum transcription may also benefit from this. Ultimately, automatic music segmentation algorithms inspired from this method may be built by exploiting the internal organization of structural segments for greater segmentation accuracy. The resulting structural metadata may then be used for music summarization (Peeters, La Burthe & Rodet, 2002), music thumbnailing (Chai & Vercoe, 2003) or interactive music playback interfaces (Goto 2006).

The S&C model also offers a valuable paradigm to build metaphoric representations of the mid-level structure of music. In the same way as this article has been abstracting musical dimensions and properties into shapes, colors or other visual features, the S&C description offers ways to represent schematically the successive mid-level elements in the musical narration and how they intuitively relate to one another at higher scales. These metadata may be of great help as an additional source of information in music teaching, since the possibility of visualizing the structure of a piece facilitates its awareness, its memorization, its understanding and its abstracted manipulation (in the composition process).

Ultimately, the S&C model may provide a complementary viewpoint to existing theories of music perception, especially regarding the concept of anticipation or of expectedness. Practically, it may contribute to computational modeling of music structure for the generation of music with a tunable level of predictability, in a way similar to (Cont 2008).

From a more technological point of view, the S&C framework points towards specific statistical models, computational approximations and data structures which appear to be well-suited to the model hypotheses, for instance :

- the use of multi-stream hidden-state models to represent S&Cs on several musical dimensions simultaneously,
- a possible simplification of the chain rule for decomposing the probability of a segment into simpler terms as a consequence of the structurally dominant dependencies within an S&C :

$$p(x_{00}x_{01}x_{10}\bar{x}_{11}) \approx p(x_{00})p(x_{01}|x_{00})p(x_{10}|x_{00})p(\bar{x}_{11}|x_{11})$$

[1] Etymologically, the italian word *cadenza* means *conclusion*.

- the benefit which architectures such as quad-trees may provide (rather than sequential models or binary trees) for representing matrix-based hierarchical relationships and dependencies between morphological elements forming an S&C,
- the use of dynamic programming techniques such as the Viterbi algorithm for optimizing the S&C description of a passage at a given time-scale.

However, this article has been carefully avoiding any explicit reference or support of these concepts in the presentation and the argumentation of the model, so as to decouple as much as possible the principles of the S&C approach from any specific computational model, implementation, architecture or algorithm.

9. SUMMARY AND CONCLUSIONS

Convergent and consistent pieces of evidence developed throughout this article tend to support the System & Contrast model as an attractive framework to characterize structural units as a set of musical elements linked to one another through some definite matrix system of properties. Relying on simple cognitive principles of logical deduction, the S&C model can indeed be applied without privileging any musical dimension nor any particular surface property of the musical content. It is therefore able to accommodate a wide diversity of musical contents.

The model is able to account for a well-defined set of elementary structural patterns which can be understood as deriving from a limited inventory of hypermetric prototypes. It is fully compatible with the latest theories of grouping (Lerdahl & Jackendoff 1983), formal functions (Caplin 1998), expectation (Narmour 1990, 1992) and cognitive rule-mapping (Narmour 2000).

The description of a music piece in terms of S&Cs at a given scale can be viewed as an optimization process which minimizes a description cost by jointly achieving selection of variables and model matching. The description cost constitutes an objective criterion to arbitrate between alternative descriptions, which can be formulated (and optimized empirically) outside of any reference to a particular computational model or algorithm.

For all these reasons, we believe that the S&C model significantly contributes towards the definition of operational concepts and methodologies for the description of structural information in music pieces, by providing :

- A neutral meta-description of the structure of musical segments which offers a means to describe accurately and consistently their inner organization, independently of their surface musical substance.
- A simple and generic scheme which is able to operate at several time-scales and to accommodate a wide variety of musical genres.
- A backward framework that encapsulates, in a single formalism, famous conventional forms (such as periods and sentences) and which appears as a compatible extension of several of the latest generative and cognitive theories of music.

In the field of MIR, the S&C model can be expected to lead to new algorithms for automatic structure extraction. It is also bound to be a base for the design of new music language models, which could partly decouple the morphological elements from the syntagmatic network of relationships, thus providing powerful generalization capabilities. Beyond MIR, the S&C model may also exhibit great potential in computer-assisted musical creation and composition, as well as in music education and teaching.

From a musicological point of view, it is perfectly clear that the S&C model does not provide an extensive and exhaustive explanation of the musical content which would preclude other levels of analysis of the musical discourse. In fact, in the same way as the grammatical analysis of a paragraph aims at explaining its internal organization in terms of the mutual relationships between words irrespectively of their meaning, the S&C model provides a framework to formulate in a standard and "neutral" way the basic logical relationships which exist between musical elements forming a structural segment at a given timescale.

The S&C analysis does *not* explain either the poïetic or the aesthetic aspects behind a musical segment. It is just useful to provide a description of the broad macro-organization of musical segments, which can be viewed as some sort of logical hypermetric <u>in relation to which</u> the analysis of specific musical dimensions can be undergone.

To what extent the principles of the S&C model are *consubstantial* to conventional music or simply the *consequence of constraints* resulting from other facets of the musical system (properties of the melodic scale, harmony rules, tonality, etc…) falls beyond the scope of the current article. But this point certainly constitutes one of the very exciting topic for further investigations of the S&C model.

Another challenging research matter is that of modeling the interaction between S&Cs at different scales, their nesting and their dependencies, which may result in an integrated view of musical structure within a whole range of time-scales.

ACKNOWLEDGEMENTS

Part of this work was achieved in the context of the QUAERO project, funded by OSEO.

REFERENCES

Abdallah S.A., Plumbley M.D. : *Information Dynamics: Patterns of expectation and surprise in the perception of music.* Connection Science, 21(2), 89-117, 2009.

Bach C.P.E. : *Essay on the True Art of Playing Keyboard Instruments.* (1753). Trans. W.J. Mitchell, p. 84. New York : W.W. Norton & Co, Inc, 1949.

Brennet M. : *Dictionnaire pratique et historique de la musique*. 1926.

Caplin W.E. : *Classical Form : A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven.* Oxford University Press, New York, 1998.

Chai W., Vercoe B. : *Music Thumbnailing via Structural Analysis*. Proc. 11th ACM International Conference on Multimedia, pp. 223-226. Berkeley, 2003.

Chomsky N. : *Syntactic Structures,* The Hague/Paris, Mouton. 1957.

Cont A. : *Modeling musical anticipation : from the time of music to the music of time*. PhD Dissertation, Univ. California, San Diego, 2008.

Cumming Naomi : *The analysis and cognition of basic melodic structures by Eugene Narmour*, Music Analysis, XI/2/3, Alexander Göhr 60th-Birthday (1992), p. 354-374.

Dannenberg R. : *Toward Automated Holistic Beat Tracking, Music Analysis And Understanding*. Proc. Annual Conference ISMIR (International Society for Music Information Retrieval Conference), pp. 366-373. London, 2005.

Deliège I. : *Prototype effect in music listening: an empirical approach of the notion of imprint*. Music Perception, Vol. 18, n° 3, pp. 371-407, 2001.

Fetis F.-J. : *La musique mise à portée de tout le monde*, Alexandre Mesnier Librairie, Paris 1830.

Goto M. : *A Chorus Section Detection Method for Musical Audio Signals and Its Application to a Music Listening Station.* IEEE Transactions on Audio, Speech, and Language Processing, Vol. 14, n° 5, pp. 1783-1794. 2006.

Grünwald P.D. : *The Minimum Description Length Principle and Reasoning under Uncertainty*. PhD Dissertation, ILLC Dissertation series 1998-03. University of Amsterdam, 1998.

Hjelmslev L. : *Essais Linguistiques* (1959, re-edited 1971, 1973 (English), 1988). Editions de Minuit.

Kolmogorov A. : *On Tables of Random Numbers.* Sankhyā, Ser. A. 25, pp. 369–375. 1963.

Lamont A., Dibben N. : *Motivic Structure and the Perception of Similarity.* Music Perception : An Interdisciplinary Journal, Vol. 18, n°. 3, pp. 245- 274, 2001.

Lerdahl F., Jackendoff R. : *A Generative Theory of Tonal Music*, MIT Press, 1983 (reprinted 1996).

Levy F. : *Complexité grammatologique et complexité aperceptive en musique.* Thèse de l'EHESS, 2003.

Liutkus A., Rafii Z., Badeau R., B. Pardo B., Richard G. : *Adaptive Filtering for Music/Voice Separation Exploiting the Repeating Musical Structure*. Proceedings IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP). Kyoto, 2012.

MacKay D.J.C. : *Information Theory, Inference, and Learning Algorithms*. Cambridge Univ. Press, 2003.

Mauch M., Noland K., Dixon S. : *Using musical structure to enhance automatic chord transcription.* Annual Conference ISMIR (International Society on Music Information Retrieval), pp. 231-236. Kobe, 2009.

Meyer L. B. : *Music, the arts and ideas: patterns and predictions in twentieth-century culture.* University of Chicago Press, 1967.

Miranda E. R. : *Composing Music with Computers*. Focal Press, 2001.

Narmour E. : *Beyond Schenkerism*. University of Chicago Press, 1977.

Narmour E. : *The analysis and cognition of basic melodic structures: the implication-realization model.* Univ. of Chicago Press, 1990.

Narmour E. : *The analysis and cognition of melodic complexity : the implication-realization model.* University of Chicago, 1992.

Narmour E. : *Music expectation by cognitive rule-mapping*. Music Perception, XVII/3 (2000), p. 329-398.

Nattiez J.-J. : *Musicologie générale et sémiologie*, Christian Bourgois Ed., 1987.

d'Ockham G. : *Summa totius logicae.* 1323.

Peeters G., Deruty E. : *Is Music Structure Annotation Multi-Dimensional ?* Proc. 3rd International LSAS (Learning Semantics of Audio Signals). pp. 75-90. Graz, 2009.

Peeters G., La Burthe A., Rodet X. : *Toward Automatic Music Audio Summary Generation from Signal Analysis*. Proc. of the International Conference on Music Information Retrieval (ISMIR), pp. 94-100. Paris, 2002.

Perone J.E. : *Form and Analysis Theory - A Bibliography*. Greenwood Press, 1998.

Rissanen J. : *Modeling by Shortest Data Description*, Automatica, Vol. 14, pp. 445-471. 1978.

Ruwet N : *Méthodes d'analyse en musicologie.* Revue Belge de Musicologie. Vol. 20, n° 1/4, pp. 65-90. 1966.

Schönberg A. : *Fundamentals of Music Composition*. Faber and Faber, London, 1967.

Smith J.B.L., Burgoyne J.A., Fujinaga I., De Roure D., Downie J.S. : *Design and Creation of a Large-Scale Database of Structural Annotations.* Proc. Annual Conference ISMIR (International Society on Music Information Retrieval), Miami (USA), 2011.

Spencer P., Temko P.M. : *A Practical Approach to the Study of Form in Music.* Prentice-Hall (Simon & Schuster), 1988 (re-edited 1994).

SELF-REFERENCES

Bimbot F., Le Blouch O., Sargent G., Vincent E. : *Decomposition into Autonomous and Comparable Blocks : A Structural Description of Music Pieces.* Proc. Annual Conference ISMIR (International Society on Music Information Retrieval), pp. 189-194. Utrecht, 2010.

Bimbot F., Deruty E., Sargent G., Vincent E. : *Methodology and Resources for The Structural Segmentation of Music Pieces into Autonomous and Comparable Blocks*. Proc. Annual Conference ISMIR (International Society on Music Information Retrieval), pp. 287-292. Miami, 2011.

Bimbot F., Deruty E., Sargent G., Vincent E. : *Semiotic Structure Labeling of Music Pieces : Concepts, Methods and Annotation Conventions*. Proc. Annual Conference ISMIR (International Society on Music Information Retrieval), pp. 235-240. Porto, 2012.

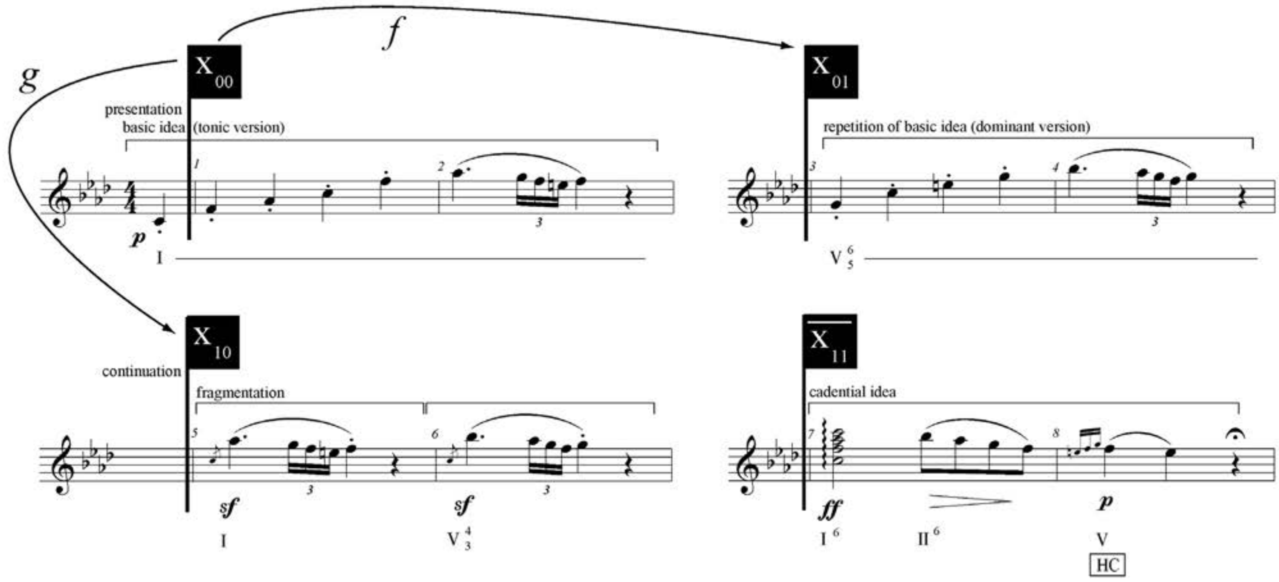
A short version of this work was formerly published as an IRISA Internal Publication, under the following reference :

Bimbot F., Deruty E., Sargent G., Vincent E. : *Complementary report to the article : "Semiotic Structure Labeling of Music Pieces : Concepts, Methods and Annotation Conventions (Proc. ISMIR 2012)"*. Publications Internes de l'IRISA - ISSN : 2102-6327. PI-1996, June 2012. 11 pages.

COMMENTED MUSICAL EXAMPLES

Example n°1

Ludwig van Beethoven – Piano Sonata in F Minor
Op. 2, 1st mvt, 1795. Bars 1-8
Adapted from (Caplin, 1998), p. 10



According to Caplin (1998, pp. 9-10), this example *"presents perhaps the most archetypical manifestation of the sentence form in the entire classical repertory"*. It starts with a *"basic idea"* that takes place in measures 1-2 and which is repeated as a dominant version in measures 3-4. The extract proceeds with what Caplin calls the *"continuation"*. Measures 5-6 show elements derived from the basic idea by means of *"fragmentation and harmonic acceleration"*, and finally, measures 7-8 present the *"cadential idea"*, which effects closure for the entire segment, using conventional harmonic and melodic formulas.

A number of properties of this sentence form can be explained in the framework of (and encapsulated within) the System & Contrast model.

- The note duration values follow an $aabc$ pattern over the passage.
- The intensity level, as indicated on the score, also matches a clear $aabc$ pattern.
- Even though the melodic lines of X_{00} and X_{01} are not an exact transposition of each other, their *direction of change* is preserved. Conversely, they greatly vary in X_{10} and $\overline{X}_{11}$. Therefore, the derivative of the melodic line is a property which forms an $a\tilde{a}bc$ pattern.
- The chord sequence tends to follow an $aba'c$ pattern[1]. It is worth noting that the rate of variation of the harmony follows an $aabc$ pattern over the passage, as it can be encoded as 0-0-1-2 in terms of number of changes within successive sub-blocks, but this seems to be a rather weak property in front of that of the harmony itself.

The relationships between measures 5 and 6 in X_{10} and measures 2 in X_{00} and 4 in X_{01} (fragmentation) are somehow overlooked by the 2x2 square S&C description, but they undoubtedly contribute to the cohesion of the segment, at a lower granularity.

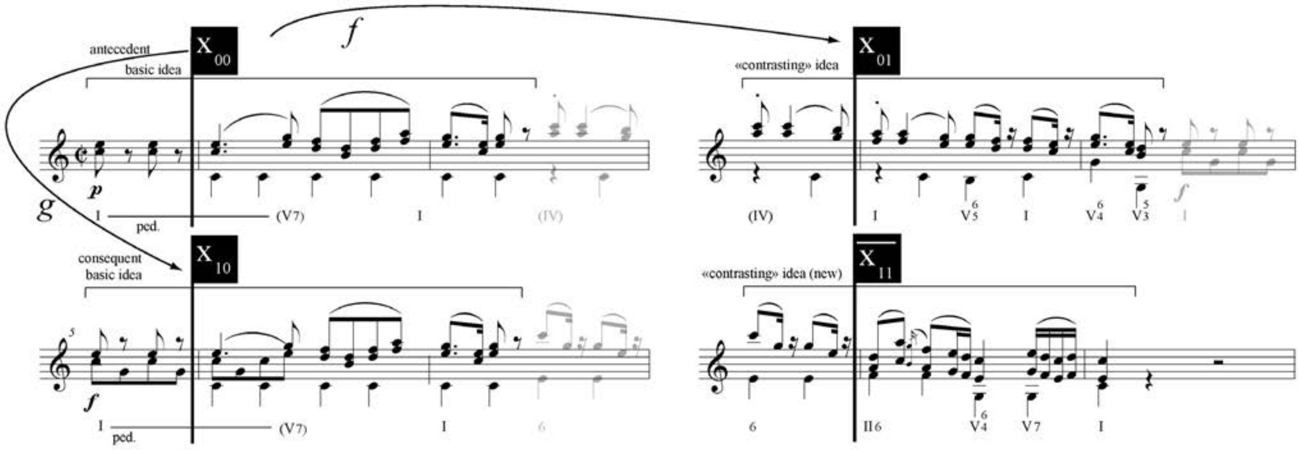
In summary, several musical dimensions in this passage follow a well-identifiable $a_1\tilde{a}_2bc$ pattern, hence the consistency of these 8 bars as a structural segment at this scale and the predominant impression towards a sentence form.

Note that it is then totally relevant, in a second step, to describe how some specific melodic formulas and harmonic cadences develop on and "embody" the structure of the segment. But its inner organization itself can neutrally be explained, with the help of the S&C paradigm, without reference to them.

[1] The a' element is also some kind of b'', where b'' would mean « ends-like » b.

Example n°2

Wolfgang Amadeus Mozart – Eine kleine Nachtmusik
Köchel 525, 2nd mvt, 1787. Bars 1-8
Adapted from (Caplin, 1998), p. 12



This example is chosen by Caplin (1998, p. 12) as an archetype of the period form. The first element is the *basic idea*. The second element is what Caplin calls a *"contrasting"* idea, which must be understood in the sense of its being "not-a-repetition" of the basic idea, and which ends, here, with a weak cadential formula. The third element is a re-exposition of the basic idea. The fourth element is an opposition to the basic idea X_{00} (but also to the “contrasting idea” X_{01}), and ends with a strong cadential formula.

Clearly, in this example, all major musical dimensions form primarily an $abac$ pattern except for minor modifications between X_{00} and X_{10}. It is also worth noting that the rhythmic pattern of a, b and c are all different but that whereas X_{00} and X_{01} are somehow off-beat in their first bar, they re-synchronize on an identical rhythmic pattern in their second bar. This is not the case for $\overline{X}_{11}$, whose melodic rhythm is in complete disparity with all the other element of the segments, thus reinforcing its contrastive status. Note also the existence of some parallelism between the harmonic placements in X_{01} and $\overline{X}_{11}$ (harmonic rhythm), which reinforces the contrastive effect of their inversion at their termination (resp. I-V versus V-I), all the more since there is obviously not such inversion between X_{00} and X_{10}.

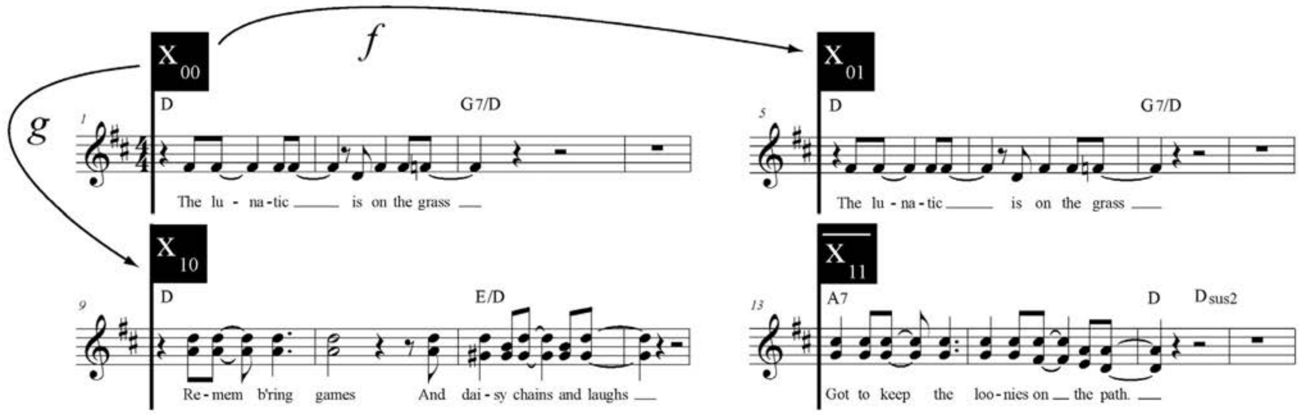
Example n°3

Johan Pachelbel – Canon and Gigue in D Major. End of XVIIth century
Bars 1-4. *Transcribed by ear into a 4 voice ensemble and transposed in C Major*

This example is based on an adapted transcription of Pachelbel's famous canon (top part). The first three measures follow what's known as the Pachelbel harmonic sequence (*marche harmonique Pachelbel)*. In such a harmonic sequence, each bar is to the preceding one what's the next bar is to the current. In other words, successive bars for a progression $a\ f(a)\ f(f(a))$ … In the top example, the fourth bar breaks away from the harmonic sequence, thus forming a contrast “IV V” instead of the non-contrastive sequel “II VI” depicted in the bottom part of the figure.

Example n°4

Pink Floyd – Brain Damage (Composer : Roger Waters)
The Dark Side of the Moon, EMI 1973. Timing : 0'15-0'43
"Pink Floyd : The Dark Side of the Moon, Guitar Tablature Edition"
pp. 109-111. Published by Music Sales America, 1992

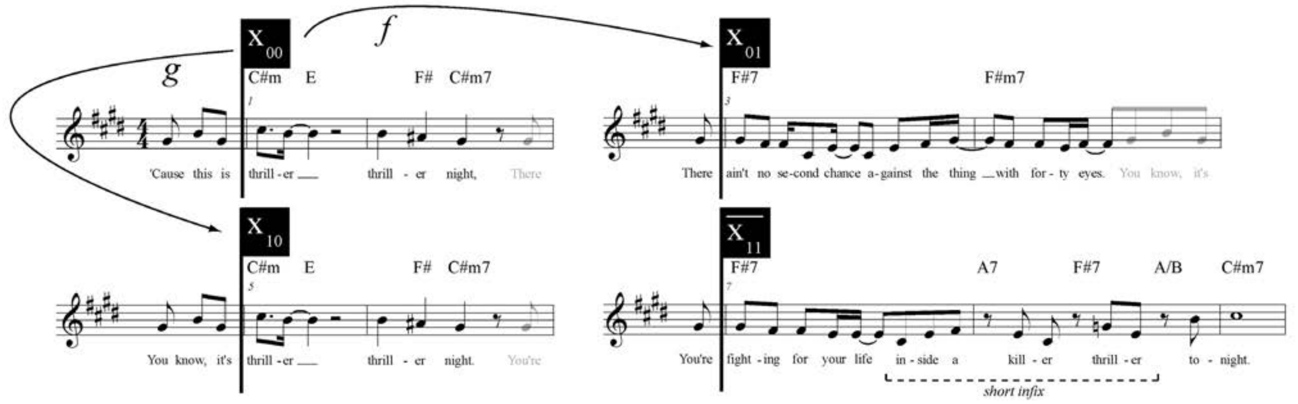


This example from a pop music piece of the 70's appears as a clear sentence structure $aabc$. While function *f* is the exact *id*, function g introduces a major reorganization of the motives X_{00} and X_{01} : the addition of a second vocal line, and denser melodic profiles repeated with an internal transposition. The contrastive segment consists in a less drastic transformation of the initial motives which shares some properties of g (the addition of a second vocal line) but where the energy distribution of the melody falls back to that of the primary elements. Therefore, the segment can be viewed as the superposition of $aabc$, $aaba'$ and $aaba$ patterns, all participating to the sentence form.

Note that here the morphological elements considered have 4 bars. It is interesting to note that sub-systems exist at the 2-bar scale, which could be primarily be considered as a plain *period* $[abab]$ followed by a sentence like form $[a_1a_2bc]$. At an even lower time-scale, a good approximation of each 4-bar group could be $[abcc][abcc][a_1b_1a_2b_2][abcc]$. This shows how the S&C model can operate at several time-scales simultaneously, thus reinforcing the consistency of the inner organization of the musical segment.

Example n°5

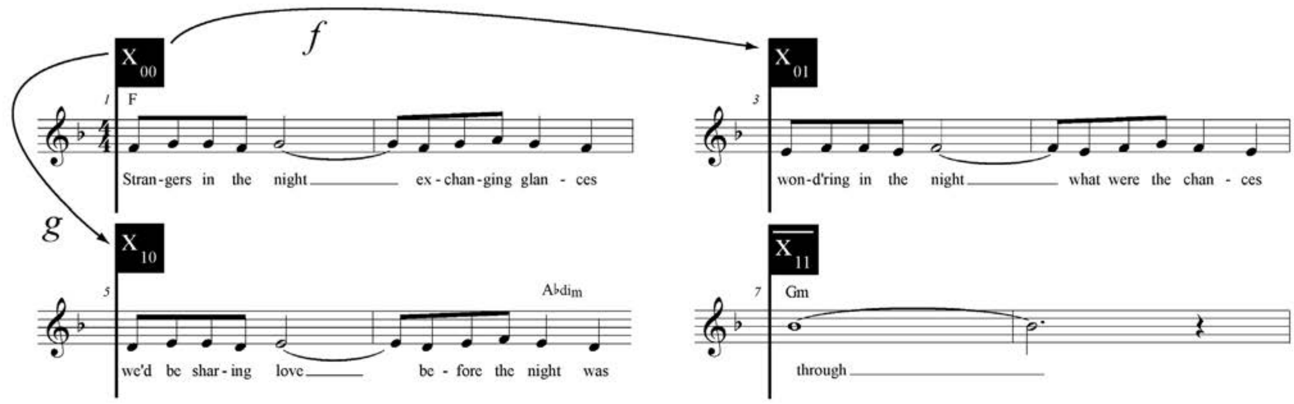
Michael Jackson – Thriller (Composer : Rod Temperton)
Thriller, EMI 1982. Timing : 2'26-2'40
"Thriller", pp. 25-26, Published by Rodsongs (PRS), 1982



This example from a famous pop song of the 80's illustrates an interesting case of *period* based on a square S&C. The function f introduces *new* musical material in X_{01}, but X_{10} presents again the material of X_{00} ($g = id$ except for the beginning of the lyrics).

The contrast function γ can be viewed as composite. On the first bar of $\overline{X}_{11}$ (numbered 7) γ is almost *id*, as this sub-segments starts very much like X_{01}. But then comes 2 bars which show a clear disparity with the end of X_{01}, namely a new, heavily syncopated motive followed by a completely steady note. Moreover, $\overline{X}_{11}$ develops over 3 bars, which also creates a contrast with the duration (2 bars) of the other morphological elements within the segment. A component of γ can thus be viewed as a stretching function creating additional contrast by the insertion of musical matter at the level of bar 8 (morphological infix) which delays the conclusion of the sequence (some sort of *phase modulation*).

Example n°6 **Frank Sinatra – Strangers in the Night (**Composers : Kaempfert, Singleton, Snyder)
Strangers in the Night, Reprise, 1966. Timing : 0’11-0’32
“Strangers in the Night”, p. 2, Universal Music Publ. Group / Hal Leonard Corp., 1966-2011

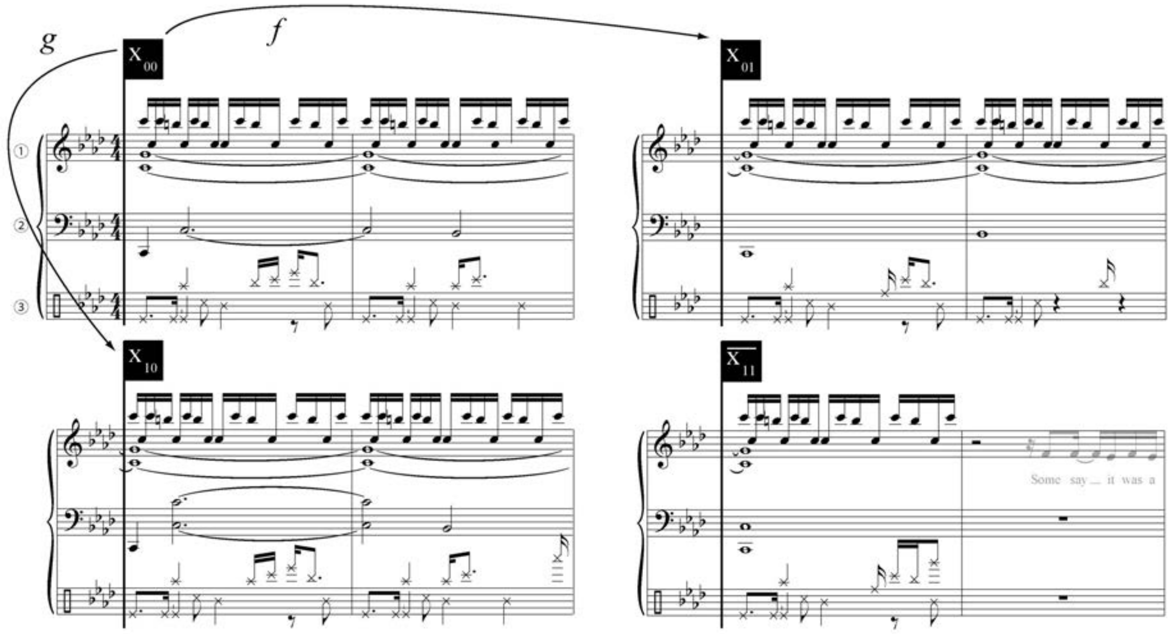


The first verse of this well-known love song of the 60’s illustrates a formal type which does not correspond either to a sentence, or to a period. It is best described as two successive diatonic transposition of the primer X_{00}, while the harmony remains constant. Obviously, the last element $\overline{X}_{11}$ introduces a completely distinct element, made of a single note that lasts seven beats and whose pitch is significantly higher than all previously heard pitches. $\overline{X}_{11}$ appears as a complete contrast with the progression installed by the three previous elements, in terms of rhythmic pattern, melodic pitch and shape, chord root and mode, musical and vocal information flow, … It can be denoted as an $[a_1 a_2 a_3 b]$ pattern, i.e. a broken progression.

Let's point out that, whereas f and $g = fof$ introduce little novelty, γ proposes something completely new. This is to be put in perspective with sentences and period, where f or g introduce much more novelty, and γ tends to be less innovative

---

Example n°7 **Nine Inch Nails – The Warning (Real World Remix)**
(Composers : Trent Reznor, Stefan Goodchild, Doudou N’Diaye Rose)
Y34RZ3R0R3M1X3D, Interscope, 2007. Timing : 0’27-0’46. *Transcribed by ear*

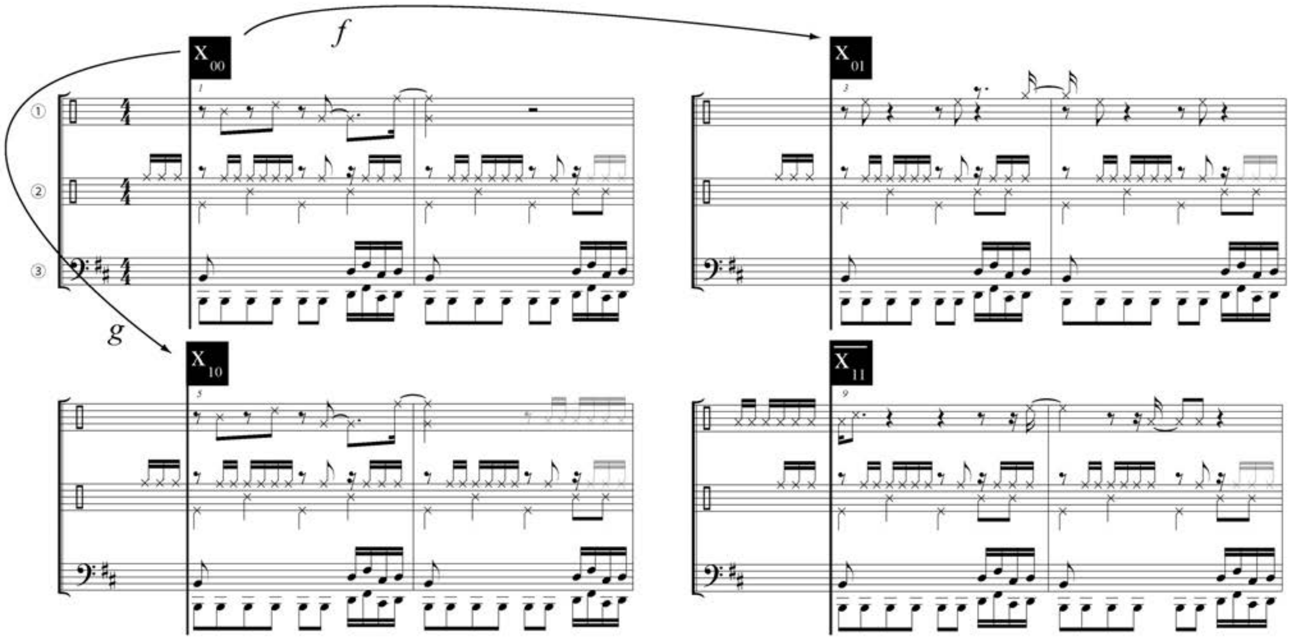


List of tracks: ① Keyboards, vocals ② Basses ③ Kick, snare, hi-hat, percussions

Whereas, in many cases, the contrast function is complex and applies diversely over the various musical dimensions, here is an instance of “industrial music” where a more radical approach can be observed : the contrast function γ simply consists in a global *suppress* function over the second part of the 4^{th} morphological element, resulting in a sudden and complete silence on all musical dimensions. This creates a definite effect of surprise, and leaves room for (and focus on) the deployment of the anacrusis of the forthcoming segment. Given the quasi-identity between X_{00} and X_{10}, the segment tends to follow a period-like behavior $abab'$.

Example n°8

Olivier Lieb – Epsilon Eridani
Epsilon Eridani EP, Bedrock Records, 2011
Timing 1'17 – 1'32. *Transcribed by ear*



List of tracks: ① Light percussive samples ② Kick, Snare & Hi-Hat ③ Basses

Considering only the pitched instruments (basses) and the traditional drum section (kick, snare, hi-hat), this segment of electronic music appears as a sequence of four identical elements, namely $aaaa$. However, careful listening reveals the presence of a set of light percussive samples organized into a period-like system $abac$. Function f matches the pattern heard in the primer X_{00} into a sequence of four syncopated, regularly spaced hits, along with another syncopated hit near the end of bar 3. The function g is the identity, and the contrast γ introduces yet another completely new pattern. In this particular case, while the conventional instruments do not convey any discriminative nor contrastive properties across the successive elements, the System & Contrast develops over non-classical musical elements in an almost inconspicuous way.

Example n°9

Britney Spears – Heaven on Earth
Composers : Mc Groarty, Huntington, Morier
Blackout, Jive, 2007. Timing : 1’23 – 1’39. *Transcribed by ear*

List of tracks: ① Lead ② Keyboards ③ Bass ④ Filtered Hi-Hat ⑤ Kick, Snare & Hi-Hat

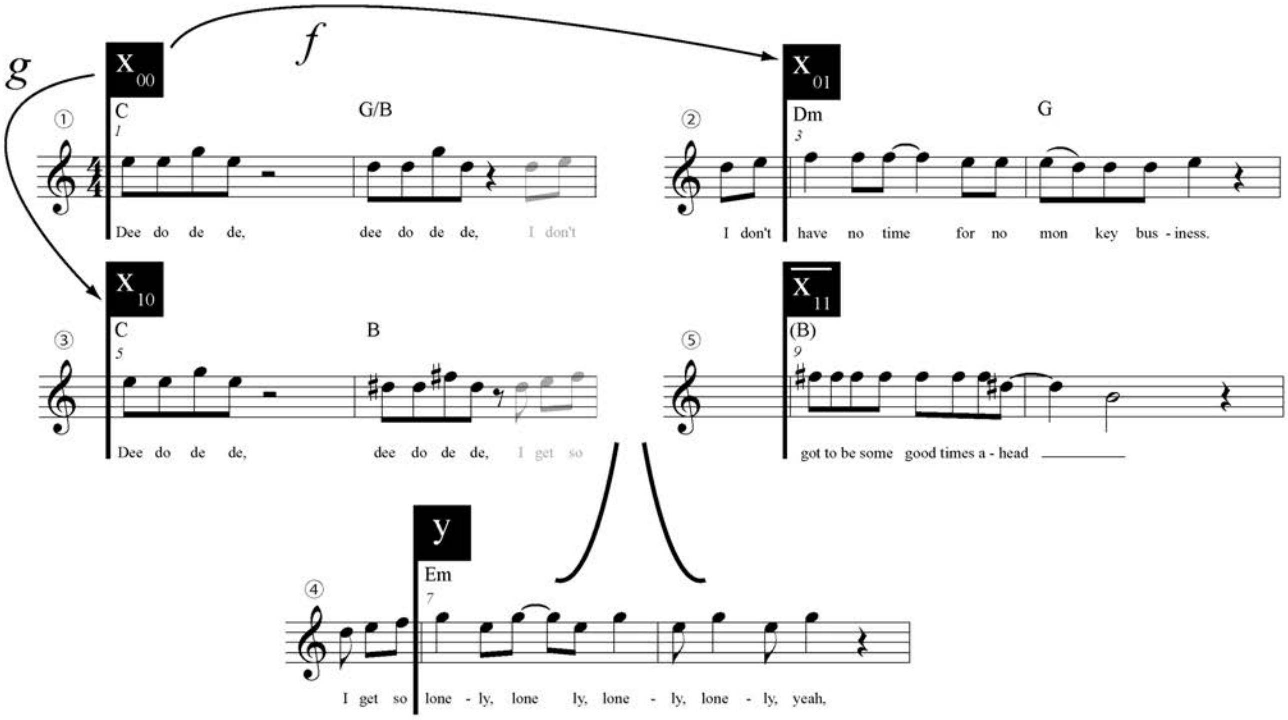
This transcribed segment from a recent American pop song exhibits a typical period like form $[abac]$ at the chosen time-scale and granularity (4 × 2 bars). Function g can be considered as “id”. Function f is a relatively straightforward transformation on the first half of the primer, but a more complex one of its second half. In fact, the sequence X_{00} X_{01} can itself be viewed as a smaller S&C at the immediately lower timescale and granularity (4 × 1 bar), where tracks 1, 3 and 4 form a contrasted system (whereas tracks 2 and 5 form a plain system). Interestingly, at an even lower time-scale (4 × 1/2 bar), track 1 of segment X_{00} exhibits a sentence-like pattern, while tracks 2, 3 and 5 show a period-like organization and track 4 a progression.

This example illustrates how the S&C model is able to account for the inner organization of musical segments at several time-scales simultaneously : the density of such multi-scale relationships within a musical segment (and especially those which involve the primer) can be considered as a positive indication of the structural consistency of that segment.

Example n°10 **Freddie Mercury – Living on my own**

Living on my Own, Parlophone, 1993. Timing : 1’22 – 1’41

“The Freddy Mercury album”, International Music Publications Ltd, 1993



This musical segment is excerpted from a song by Freddie Mercury originally released in 1985 in the album “Mr Bad Guy” and remixed in 1993 as a dance version, under which it became famous. This example illustrates a pentadic system, denoted here as $\mathrm{X}_{00}\,\mathrm{X}_{01}\,\mathrm{X}_{10}\,\mathrm{Y}\,\overline{\mathrm{X}}_{11}$, i.e. as a square S&C stem with an extraneous infix.

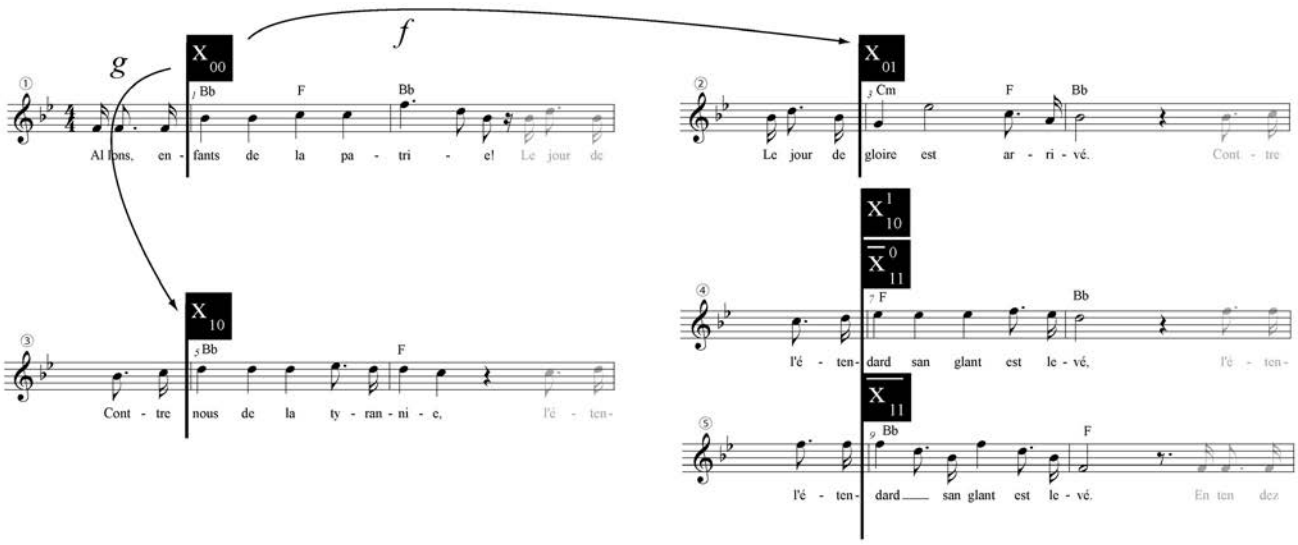
After the exposition of the primer X_{00}, function f formulates the introduction a new element X_{01}. Function g relates the primer to a third element X_{10} which *starts-(a)like*, thus inducing a period-like pattern. The first bars in both elements are indeed identical, whereas the second bars can be deduced from each other by a chromatic transposition, creating a tonality shift acting as a transition towards the following element.

The fourth element (labeled as Y) could be the contrast of the system, with repetitions, syncopated patterns and a steady tonality of E minor, none of these properties being observed in the previous elements. However, the fifth element, $\overline{\mathrm{X}}_{11}$, also appears to relate to the system : it constitutes somehow a return to non-syncopated rhythmic patterns, its tonality is in line with that of X_{10}. Moreover, it cannot serve as a primer for the forthcoming passage but it can stand as a valid contrast in a square stem $\mathrm{X}_{00}\,\mathrm{X}_{01}\,\mathrm{X}_{10}\,\overline{\mathrm{X}}_{11}$, from which Y would be removed.

According to this description, the inner structure of the segment can thus be written as $aba'yc$, with c representing the contrast and y an additional extraneous element (based on an “*insert*” function h). At this point, it is worthwhile noting that the infix Y is actually reused, in a different context, at the end of another song of the same album (“I was born to love you”, ~ 4’30”), which further supports the hypothesis of an exogenous (or at least separable) 4th element.

Note however that an alternative description could consider that bars 7-10 form the contrast of the segment, namely a double-size contrast with respect to the size of the carrier system elements (which itself constitutes a factor of contrast).

Example n°11 **Claude-Joseph Rouget de Lisle – La Marseillaise (National French Anthem)**
Adapted from l'« Hymne des Marseillais », 1792. Bars 1-10. *Transcribed by ear*



This passage is another example of a pentadic system, which is here written as $X_{00}\ X_{01}\ X_{10}\ \begin{matrix} X_{10}^{1} \\ \overline{X}_{11}^{0} \end{matrix}\ \overline{X}_{11}$.

The 4th element can indeed be viewed as forming (i) partly a sub-system with the 3rd element X_{10} and (ii) partly a sub-system with the 5th element $\overline{X}_{11}$, depending on the musical dimensions. The following patterns can be observed :

rhymes	$ababb$
harmony	$abcc^{*}c$
melodic contour	$aba'a'c$
rhythm of the melodic lead	$aa'bb'c$

As a consequence, the 5 elements happen to form a consistent S&C, where the 4th element relates equivocally to both its neighbors. However, the sequence $X_{00}\ X_{01}\ X_{10}\ \overline{X}_{11}$ seems to form a valid square stem (and element $\overline{X}_{11}$ is not eligible as a valid primer for what comes next in the piece). Therefore, the segment can be considered as a pentadic system with a dominating 2×2 & 1 shape. Here again, an alternative would be to consider that the last 4 bars form the contrast, on the basis of an oversized last element.

Example n°12

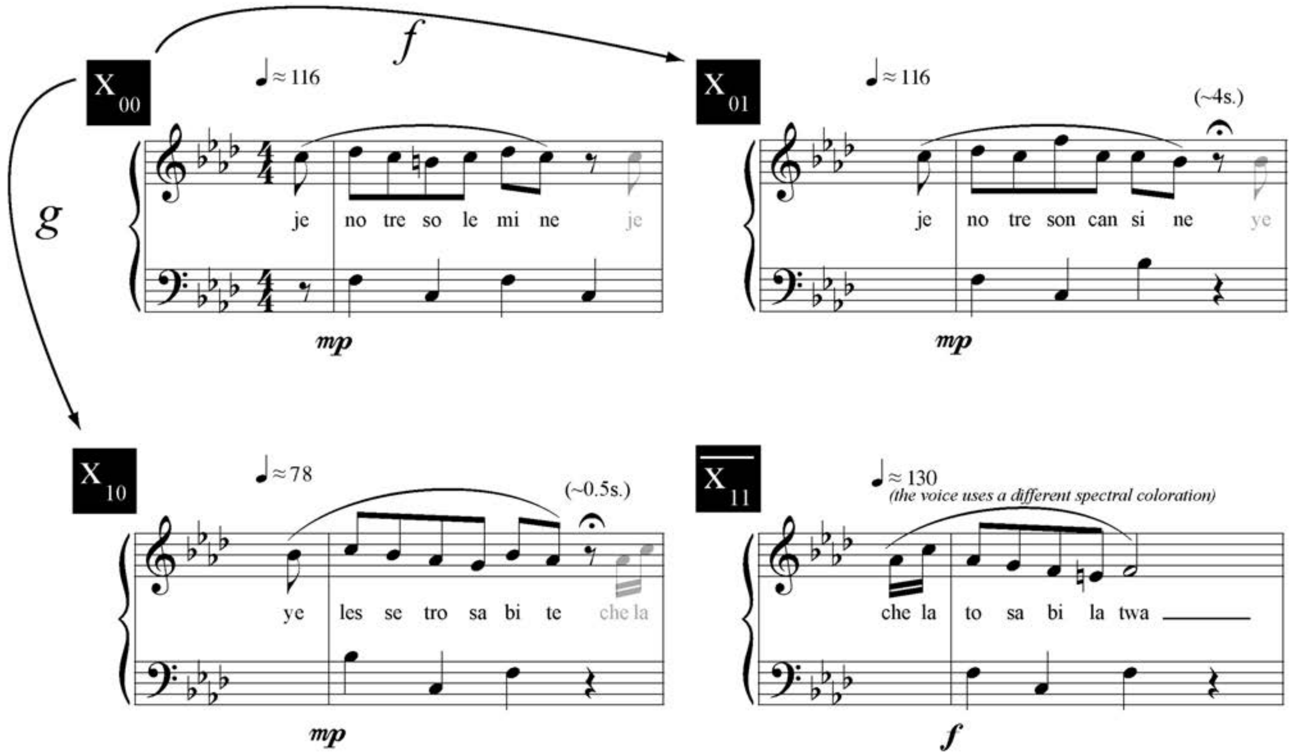
Antonio Vivaldi – Concerto n°3 in F Major ("Autumn")
Op. 8, Ryom Verzeichnis 293, 1st mvt, 1723. Bars 1-6
Reduction from a transcription by H. Sawano (http://sound.jp/kazane)

These first 6 bars of the first movement of the "Autumn" concerto by Vivaldi is a simple example of an (almost) plain "wide" hexadic system. Such a system can be formalized using a primer and 3 syntagmatic functions f_1, f_2 and g. Here, $f_1 = id$, whereas $f_2 \simeq new$. Function g combines a change of dynamics and the transposition of three of the four voices to the lower octave. Here, the contrast γ is the $\widetilde{id}$ function, i.e. identity, except for the last part of the accompaniment of X_{12}, where it can be seen that (on this transcription), function g stops applying to one of the voices. This is equivalent to assuming that γ is equal to id in the first part of X_{12} and then turns into g^{-1} for that particular voice, towards the second part.

Example n°13 **Charles Chaplin – Untitled Song from the "Modern Times" movie (1936)**

After Leo Daniderff – Je cherche après Titine (1917)

Timing : 1:18'28-1:18'43. *Transcribed by ear*



Each verse of this famous nonsensical version of a French song dating from 1917, and which went around the world in the late 30's with Chaplin's movie, is made up of four 1-bar elements.

Among the most salient properties forming the S&C :

- the rhythmic pattern of the melody goes $aaac$,
- the underlying harmonic support of the segment goes abb^*c, where $a =$ Fm Fm, $b =$ Fm Bm, $b^* =$ Bm Fm the reverse chord progression of b, and $c =$ C Fm a clear contrast with any simple system that could be imagined to explain a carrier system based on a, b and b^* (most probably abb^*a, since $a^* = a$).

The melodic line forms almost an $aba'c$ but it varies slightly from one verse to another in Chaplin's version.

In Chaplin's interpretation, the fourth element recurrently creates laughter and/or applause from the audience, in reaction to the intended surprise effect. Central to the acting performance of Chaplin is the reinforcement of the disparity of the 4th musical element $\overline{X}_{11}$ by creating additional contrasts in terms of dynamics $[mp\ mp\ mp\ f]$, tempo $[116\ \ 116\ \ 78\ \ 130]$, rhyme of the (fake) lyrics $[e\ e\ e\ wa]$, voice timbre modification on the last element, etc…

Example n°14 **Richard Wagner – Gurnemanz' Monologue**

Parsifal (Act I). Wagner Werk Verzeichnis 111, 1882

"Parsifal, Klavierauszug zu zwei Händen", Ed. R. Kleinmidjel, p. 19, 1911

This excerpt from Wagner's Parsifal is an example of a musical passage which cannot be described (at any conceivable mid-level time-scale) using the System & Contrast model. It is indeed striking how, in this musical flow, it is impossible to identify elements that relate to each other on any musical dimension. If this example may be viewed as an instance of "infinite (or "unending") melody, it can certainly be understood as music based on permanent novelty, as it seems designed so as to exhibit very little redundancy at the morpho-syntagmatic level.

Example n°15

Wolfgang Amadeus Mozart – Piano Sonata in B-Flat
Köchel 333/315c, 3[rd] mvt, 1783. Bars 55-103
Annotated screen capture from (Caplin, 1998), p. 240

EXAMPLE 16.8 Mozart, Piano Sonata in B-flat, K. 333/315c, iii, 55–102

In this example, we annotate an entire passage from Mozart's piano sonata in B-Flat, K333/315c, third movement into successive S&Cs. The example is borrowed from (Caplin, 1998) - p. 240, who uses it as an illustration for a particular part of the *sonata-rondo* form where the "couplet #2" is used as a "double-region couplet", i.e. a couplet featuring two main themes.

For this passage, we present a comprehensive description of bars 57-102 in terms of S&Cs. We work at a granularity of 2 bars and we identify 6 successive S&C, denoted A-F, whose main properties are given in the table below.

Code	Description	Shape
A	A standard square S&C dominated by a sentence structure	$2 \times 2\,.$
B	A standard square S&C dominated by a sentence structure	$2 \times 2\,.$
C	A shorter square S&C at a half time-scale (acting as a fast transition) Segment C is tiled with segment D over 1 bar	$\frac{1}{2}(2 \times 2)\,.\,(-\frac{1}{2}+)$
D	A sophisticated 8-element segment whose inner structure can be decomposed as $[\boldsymbol{a_1}\ \boldsymbol{a_2}\ \boldsymbol{b}\ b\ y_1\ y'_2\ c^0\ \boldsymbol{c}]$, i.e. a period-like square stem (elements in bold font) and 4 affixes : 1 redundant, 2 nested and 1 pre-contrast.	$(2 \times 2)\ \&\ 1\ \&\ \frac{1}{2}(2 \times 2) + 1\ .$
E	Segment D is tiled with segment E over 1 bar Segment E is a quasi-plain S&C exhibiting a period-like pattern $a_1 b_1 \tilde{a}_2 b_2$	$\left(-\frac{1}{2}+\right) 2 \times 2\ :$
F	This set of 4 bars can be viewed as a half-timescale S&C $a_1 a_2 a_3 b$. It acts as a dyad which follows a plain square system. The sequence EF forms a triadic S&C $[a_1 a_2 b]$ at the immediately upper time-scale	$\frac{1}{2}(2 \times 2)\,.$
		23 morphological units i.e. 46 bars

According to Caplin, theme #2 (which corresponds to segment D) is described as a *"loosely constructed sentential unit"* including a standard *presentation* during measures 76 to 79, followed by a complex *continuation* that lasts until measure 89. By describing accurately the “syntax” of this construction as $[a_1\ a_2\ b\ b\ y_1\ y'_2\ c^0\ c]$, the S&C model offers much more insight on the inner organization of the segment, identifying different phases of the *continuation* and their relationship with the stem of the segment.

Example n°16

Loreen – Euphoria
Composers : T. G:son & P. Boström. Label : Warner Music
Eurovision Song Contest winner 2012 (Sweden). *Transcribed by ear*

On the above transcription of the first 1’20” of this trance-inspired euro-pop song, we consider structural segments corresponding to 8 bars (i.e. one line) and each morphological element lasts 2 bars.

While the inner structure of each of the first two verses (bars 1-8 and 9-16) clearly forms a dominant period-like pattern $abac$ (supported by a plain $abab$ rhyme system), the chorus proceeds to a sudden change of *regime* : the harmony becomes $abab$, the melodic system exhibits a dominant $abb'c$ pattern[1], prone to create a structural uncertainty on the status of bars 17-18 (“Euphoria”) as a primer.

This ambiguity is however resolved by at least two other factors :

- Bars 23-24 clearly create a contrast with the 3 previous morphological elements (17-22), forming a very distinctive “b” in an $aaab$ pattern in terms of (i) “on-the-beat” vs “off-beat” note placement, and (ii) “rising” vs “falling” melodic termination. In particular, the insistent and syncopated repetition of “up, up, up…” over 1½ bar constitutes a definite “punctuation mark” in this context, in contrast with the smoother flow of the previous bars.
- Hypothesizing a structural segment which would start on bar 19 (rather than 17) would create a very inconsistent organization for the *second part* of the chorus (4th line of the transcription), given that, bars 33-34 necessarily constitute the primer of a new occurrence of the song’s verse (5th line).

This example illustrates how competing hypotheses for isolated segments can be arbitrated by more global considerations taking into account successive segments.

[1] Exactly $abbc$ for what concerns the rhymes