

**Rapport final de la partie « *Modélisation* » du projet  
« *Segmentation et reconnaissance de patterns dans la musique :  
Perception, analyse, modélisation.* »**

## **Résumé**

La perception structurelle sous-jacente à l'écoute musicale est un processus cognitif d'une grande complexité qui met en jeu une multitude de stratégies interdépendantes. La compréhension approfondie de cette architecture cognitive, et en parallèle la description détaillée du foisonnement de structures émergentes, nécessitent en conséquence le support de la systématisation informatique, afin d'appuyer la démarche empirique sur une modélisation complexe. Les multiples stratégies concourant à l'écoute musicale structurelle peuvent être scindées en diverses catégories. Une distinction majeure, développée dans la collaboration entre l'équipe expérimentation et l'équipe modélisation, a été établie entre d'une part les mécanismes d'émergence structurelle de bas niveau apparentés à la *Gestalt*-theorie, et d'autre part les stratégies de plus haut niveau fondées sur des schémas culturels.

Le versant modélisation de la collaboration a été axé dans un premier temps sur les mécanismes de bas niveau, notamment la problématique d'extraction de motifs répétés au sein de partitions musicales. Un espace multi-dimensionnel de paramètres musicaux a été défini, fondé sur des descriptions de surface réparties suivant les composantes mélodiques (hauteur diatonique et chromatique des notes, distance intervallique entre notes successives, sens de variation du contour mélodique) et rythmiques (valeur rythmique, positionnement des notes par rapport à la pulsation métrique). La recherche de motifs s'effectue de manière adaptative au sein de cet espace. Nous avons montré en particulier l'intérêt musical, et la faisabilité algorithmique, d'une construction de motifs dits *hétérogènes*, composés d'une succession de descriptions définies sur un ensemble *variable* de dimensions musicales. Ceci permet de rendre compte de manière plus précise des structures motiviques caractéristiques des œuvres musicales, qui peuvent parfois présenter des particularités locales au niveau de certaines dimensions musicales uniquement, tel qu'un motif mélodico-rythmique dans une première phase, puis simplement rythmique dans une seconde phase.

Nous avons mis en évidence les difficultés sous-jacentes à la problématique d'extraction automatique de motifs répétés, et souligné les insuffisances des approches musicologiques, cognitives ou informatiques antérieures. Les principes de base caractérisant ces mécanismes de bas niveau sont modélisés d'une manière la plus générique possible, afin d'éviter une spécialisation précoce du modèle vers un champ restreint de possibilités. La problématique majeure émergeant d'une telle modélisation cognitive est relative à l'explosion combinatoire de structures musicales admissibles. L'approche proposée consiste en une explicitation des facteurs de redondance structurelle sous-jacents à une telle problématique, et en l'élaboration d'heuristiques de contrôle, assurant une meilleure concordance du comportement algorithmique avec les attentes musicologiques et les intuitions perceptives. Ainsi, une importante part de la redondance combinatoire est maîtrisée avec l'aide de la théorie

mathématique des concepts formels, ainsi que son pendant informatique de motifs fermés. Ces heuristiques permettent de filtrer de manière maximale les descriptions redondantes – consistant en des motifs moins spécifiques que d'autres motifs déjà découverts – tout en évitant toute perte d'information structurelle. Un motif moins spécifique – par exemple, un suffixe d'un motif donné, ou encore la composante strictement rythmique d'un motif mélodico-rythmique – est en général une structure redondante qui n'apporte pas d'information supplémentaire. Cette description moins spécifique sera toutefois conservée si elle apparaît occasionnellement sous forme d'occurrences autonomes au sein du document musical, c'est-à-dire non intégrées à des occurrences du motif plus spécifique. Nous avons en outre développé un nouvel appareil conceptuel centré sur la notion de motifs cycliques afin de contrôler un autre facteur important de redondance combinatoire, qui apparaît lorsqu'un motif est répété plusieurs fois de manière successive, induisant une séquence périodique. De telles répétitions périodiques sont monnaie courante dans le discours musical : trilles, lignes d'arpèges, répétition prolongée d'une cellule rythmique ou mélodique, voire tout suite de valeurs rythmiques constantes. Une telle configuration engendre une combinatoire importante de structures motiviques redondantes. L'introduction du concept de motif cyclique permet de contrôler de manière adaptative et élégante ce facteur de redondance. D'autres comportements algorithmiques problématiques ont été observés, notamment des formes alternatives de redondance combinatoire, et ont été maîtrisés par l'intermédiaire d'heuristiques supplémentaires, fondées soit sur des principes cognitifs connus (tels que des règles de Gestalt), ou sur de nouveaux principes originaux qui pourront être approfondis par la suite. La modélisation résultante, offre une description à la fois compacte et détaillée des documents musicaux, présente un comportement surpassant significativement l'état de l'art actuel, et peut donc avoir des applications intéressantes en musicologie. D'autre part, en raison du haut degré de stabilité du modèle et de pertinence des résultats, la véracité cognitive de la modélisation pourra être envisagée dans des travaux ultérieurs.

La seconde phase du travail de modélisation est consacrée à l'articulation entre mécanismes de bas niveau et stratégies fondées sur des schémas culturels. La modélisation constituée lors de la première phase, bien qu'orientée vers les mécanismes de bas niveau, est implicitement influencée par le contexte culturel, lequel s'exprime à travers les corpus musicaux expérimentés et les intuitions mises en oeuvre au cours de la démarche de modélisation. Une confrontation du modèle avec des corpus musicaux extra-européens permet un remodelage des mécanismes de bas-niveau et une prise en compte des variantes interculturelles en termes de dépendance à des schémas culturels. L'intégration de schémas culturels au sein de la modélisation est envisagée sous deux formes : dans un premier temps, des schémas canoniques offerts par l'expertise ethno-musicologique, puis dans un second temps, des schémas résultants des analyses statistiques des tests d'écoute. Cette seconde phase de notre collaboration se poursuit actuellement en prolongement de l'ACI proprement dite.

## **1. Préambule**

Le projet de modélisation informatique engagé dans ce projet a pour objectif principal d'offrir à la musicologie – et notamment à l'éthnomusicologie – de nouveaux outils permettant une description approfondie des configurations structurelles présentes au sein de chaque œuvre musicale. L'apport de l'informatique est multiple : l'automatisation du processus d'analyse permet en particulier un approfondissement de l'analyse. Cette automatisation implique en outre une formalisation totalement explicite des mécanismes mis en oeuvre lors de ces procédures d'analyse, formalisation qui

pose de nombreuses difficultés, comme nous allons le constater, et qui, en dépit des nombreuses recherches sur ce sujet (Cf. par exemple [CAMBOUROPOULOS 2006], [COPE 1996], [MEREDITH et al 2002], [ROLLAND 1999], [ROWE 1993], [TEMPERLEY 1988]), n'a pas encore été véritablement achevée. Les processus de découverte de structures musicales présentent en effet une grande complexité, qui nécessite une modélisation la plus soignée et détaillée possible des stratégies cognitives mises en œuvre lors de l'écoute musicale. Il convient en outre de distinguer, au sein de ces stratégies, ce qui relève d'une acculturation à un milieu culturel particulier, de ce qui se fonde au contraire sur des mécanismes de perception élémentaires et innés. La perception de structures musicales est ainsi régie par un certain nombre de stratégies indépendantes en interaction : certaines de ces stratégies sont directement issues de schémas cognitifs culturels, d'autres proviennent de mécanismes élémentaires de détection de groupements de notes. D'où l'intérêt de mener cette tâche de modélisation en collaboration étroite avec les sciences cognitives et l'éthnomusicologie [McADAMS et al. 2004], collaboration qui a pu être initiée dans le cadre de l'A.C.I. « Systèmes complexes en sciences humaines et sociales » financée par le CNRS.

Parmi les mécanismes de découverte qui ne relèvent pas directement de schémas cognitifs culturels peuvent être distinguées deux heuristiques particulières. D'une part, une séquence musicale peut être décomposée en une succession de segments, chaque segmentation étant provoquée par une discontinuité suivant un ou plusieurs attributs auditifs, telle qu'un saut brusque de hauteur, une longue durée ou un silence, etc. (Cf., par exemple, [TENNEY et POLANSKY 1980], [LERDAHL et JACKENDOFF 1983], [TEMPERLEY 1988], [NARMOUR 1990], [CAMBOUROPOULOS 2006]). D'autres structures musicales, par contre, émergent en raison de leur répétition au sein de la pièce musicale [RUWET 1966]. Dans notre approche, nous nous limiterons dans un premier temps à cette simple recherche de répétition, car, comme nous allons le mettre en évidence dans ce rapport, une telle problématique, malgré son apparente simplicité, soulève de nombreuses difficultés. En effet, les tentatives musicologiques de description motivique détaillée (Cf. par exemple [RETI 1951], [MEYER 1956], [RUWET 1966], [NATTIEZ 1975]) ont rencontré des difficultés à démontrer la pertinence des motifs découverts, et se voient également dans l'incapacité d'assurer une exhaustivité de l'analyse. C'est ici que la modélisation informatique peut apporter une aide précieuse, en permettant une description précise de l'ensemble des heuristiques de découverte mises en œuvre, et en assurant une mise en application exhaustive de ces mécanismes [LARTILLOT et SAINT-JAMES 2005].

Nous allons justifier dans un premier temps la nécessité d'une modélisation informatique de stratégies d'écoute à travers une brève revue critique des réflexions épistémologiques concernant l'analyse motivique. Nous montrerons ensuite pourquoi les tentatives informatiques d'automatisation de cette recherche de répétitions ne parviennent pas à offrir des résultats qui concordent avec les attentes musicologiques. Nous mettrons en évidence en particulier l'existence d'une problématique de redondance combinatoire, qui provient de la définition-même de la tâche de recherche de répétitions

motiviques. On pose pour hypothèse qu'une telle redondance est implicitement gérée au sein du système cognitif qui fonde aussi bien l'écoute musicale que l'intuition du musicologue. Une reconstitution de ces mécanismes de gestion de la redondance permettrait alors d'améliorer la pertinence perceptive des résultats. Certaines de ces heuristiques – telles que le concept de motifs clos – s'apparentent directement aux travaux actuels en algorithmique d'extraction de connaissance ; d'autres – telles que la règle de *Gestalt* de figure/fond – offrent une modélisation concrète de phénomènes étudiés dans les sciences cognitives ; d'autres enfin ouvrent de nouvelles problématiques encore peu étudiées jusqu'alors, telles que la gestion du réseau de spécificité.

La mise en œuvre informatique d'une telle modélisation d'inspiration cognitive donne lieu effectivement à des résultats surpassant, à certains égards, les approches informatiques contemporaines, notamment en raison de la capacité à offrir une description assez claire et pertinente de structures motiviques. Ce modèle n'est toutefois qu'une première ébauche d'un système complet d'analyse motivique, car il ne peut analyser actuellement que des oeuvres musicales suffisamment simples, et ne peut assurer non plus une pertinence totale des résultats. De nombreux mécanismes doivent être pris en compte afin d'intégrer l'ensemble des transformations possibles et d'envisager l'expression musicale dans toute sa généralité polyphonique.

Dans le cadre de l'A.C.I., l'application de cette méthodologie à l'analyse d'improvisations monodiques modales a nécessité la prise en compte de certaines transformations musicales typiques de la musique modale tunisienne et notamment le processus d'ornementations musicales, au cours duquel diverses répétitions d'un même motif originel sont agrémentées de notes supplémentaires décoratives. Nous proposons une méthode de résolution du problème qui permet un filtrage adaptatif des ornements sans perte d'information. Cette méthode consiste d'une part en la construction d'un graphe, appelé graphe syntagmatique, mettant en évidence les relations de successivités temporelles au sein des notes constituant la monodie improvisée. La structure profonde que constitue le motif originel peut ainsi être retrouvé à partir de la surface ornementée suivant un parcours sélectif de certaines relations de ce graphe syntagmatique. Les résultats de la modélisation computationnelle sont par la suite traduits sous forme d'une segmentation hypothétique de l'improvisation musicale. Ceci permet une comparaison des résultats offerts par la modélisation computationnelle avec ceux produits par les tests d'écoute. La modélisation peut par la suite être enrichie, suivant une démarche de réduction progressive des divergences entre les résultats offerts par les deux versants de la collaboration.

## **2. Enjeux et difficultés d'une modélisation informatique.**

### **2.1. Automatiser afin d'approfondir l'analyse musicale.**

La musique, domaine d'expression aux potentiels de créativité infinis, se décline en une multitude de productions singulières. Il se constitue ainsi un corpus en perpétuel accroissement, dont il

n'est pas aisé de produire une vision synthétique. De plus, chaque oeuvre musicale offre une structuration d'une grande richesse. Une telle complexité structurelle résulte d'un acte de création, qui n'est en général pas contrôlé de manière totalement explicite par le compositeur, et encore moins par l'improvisateur : les multiples processus mis en jeu lors du geste créateur jaillissent de manière intuitive et spontanée, sans réelle possibilité de contrôle rationnel total et de formalisation explicite. De même, l'auditeur ne perçoit généralement d'une oeuvre musicale qu'une partie de sa grande complexité structurelle, et ne peut rendre compte de manière explicite qu'une fraction encore plus réduite. C'est une des tâches essentielles de la musicologie d'entreprendre une compréhension et une explicitation de la complexité du phénomène musical par l'intermédiaire de formalisations et de schématisations. La théorie musicale peut guider la démarche analytique, qui se divise ainsi en autant de composantes que de dimensions offertes par la théorie. Malgré tout, l'oeuvre musicale, ne se réduit pas à une simple adéquation à la théorie musicale, mais conserve une part de singularité propre, qui constitue sa richesse essentielle.

Nous nous intéresserons en particulier à l'analyse motivique, qui met en évidence, au sein de chaque oeuvre musicale, un dense réseau de motifs musicaux en étroite interrelation. Rudolf RETI (1951) proposa une démarche d'analyse thématique permettant de dépasser les approches musicales traditionnelles généralement réductionnistes en mettant en évidence la richesse des développements thématiques élaborés au cours d'une oeuvre musicale. Ses analyses étaient cependant focalisées sur un seul aspect de ce développement, car elles étaient fondées sur une vision particulière de l'oeuvre musicale, empreinte de l'esthétique propre à Arnold Schönberg, comme entièrement bâtie à partir d'un seul motif germinal. Et même dans le cadre d'un champ d'étude aussi limité, l'analyse peut être difficilement effectuée de manière exhaustive, en raison de la grande complexité structurelle des développements motiviques. Les méthodes d'analyses musicologiques ont pour principal intérêt de rendre explicites les structurations perçues ainsi que les stratégies d'écoute mises en oeuvre. Mais la réalisation manuelle de ces techniques analytiques nécessite un travail fastidieux et rébarbatif, qui peut parfois faire oublier les enjeux perceptifs sous-tendant initialement une telle tâche. Ainsi l'analyse thématique de Rudolf Reti, selon Nicholas COOK (1987), « dégénère en un exercice purement mécanique suivant lequel la partition est analysée sans être véritablement lue »<sup>1</sup>. L'informatique peut ici répondre à une telle limitation, car l'automatisation des outils de découverte automatique de structures musicales rend possible une étude exhaustive du problème. Là où les approches manuelles doivent limiter, pour des raisons pratiques, la complexité des mécanismes étudiés<sup>2</sup>, le support informatique permet une étude approfondie de la complexité musicale. Ainsi, l'informatique permet de

---

<sup>1</sup> p. 114.

<sup>2</sup> Par exemple, la structure de groupement proposée par Lerdahl et Jackendoff (1983), qui a certes donné lieu à une implémentation informatique (Temperley 1988), se fonde cependant sur un nombre limité de règles simples, et sur une structuration strictement hiérarchique des motifs musicaux, ce qui limite la compréhension de la structure musicale.

radicaliser la créativité des procédures de découverte, d'élargir le champ des analyses, aussi bien au niveau intensif du détail des analyses, que celui extensif de la taille du corpus à analyser.

## **2.2. Modéliser afin d'explicitier les mécanismes de découverte.**

Une autre difficulté de l'analyse motivique, et de l'analyse thématique de Rudolf Reti en particulier, concerne la justification des heuristiques adoptées. Nicholas Cook, par exemple, a mis en évidence la part importante de subjectivité des analyses de Reti, et a remis en cause en particulier la pertinence objective des stratégies choisies. Dans le sillon de la formalisation linguistique ont été proposées des tentatives de modélisation explicite des stratégies de découverte de structures musicales. La proposition qui a vraisemblablement le plus marqué la musicologie du vingtième siècle s'intègre dans un projet de sémiologie musicale (NATTIEZ 1975) et se fonde sur une méthode dite d'*analyse paradigmatique*, initiée par les travaux musicologiques du linguiste Nicolas RUWET (1966), lesquels prolongent et systématisent certains principes élaborés par l'éthnomusicologue Gilbert ROUGET (1961) et reprend une technique d'alignement de séquence proposée par Claude LÉVI-STRAUSS (1958). Ruwet a manifesté la nécessité d'une formalisation rigoureuse et explicite des procédures de découverte de structures musicales, procédures relevant alors d'une approche dite *analytique*. À cette approche s'oppose la démarche *synthétique*, qui consiste en une re-découverte, ou reconnaissance, de schémas musicaux déjà connus. L'auteur défend donc l'urgente nécessité d'une méthodologie analytique entièrement formalisée :

Les analyses musicales, même les meilleures — par exemple celle donnée par Pierre Boulez du *Sacre du Printemps* — ne formulent pas les critères de découverte sur lesquels elles reposent. D'une manière générale, la plupart des traités d'harmonie, de fugue, etc., présentent une situation analogue à celle offerte par les grammaires traditionnelles : le modèle est synthétique, partiellement explicite seulement, et entaché de normativisme.<sup>3</sup>

L'approche de Ruwet se concentre sur la recherche de répétitions de motifs. Elle se fonde sur un certain nombre d'hypothèses concernant la structure musicale : en particulier, il est supposé que la structure de chaque œuvre musicale est de nature taxinomique : elle se compose d'une hiérarchie stricte de structures à plusieurs niveaux emboîtés les uns dans les autres. Mais la généralité d'une telle hypothèse peut être discutée, notamment dans un contexte éthnomusicologique, car elle ne semble pas pouvoir être vérifiée de manière universelle. Quoi qu'il en soit, si une telle hypothèse est valable, la découverte effective de cette structuration hiérarchique s'effectuerait en prolongement des préceptes de la linguistique structuraliste (Cf. HJELMSLEV, 1966), de deux manières :

soit « de bas en haut », par une détermination des unités élémentaires puis un assemblage récursif des unités en unités de niveau supérieur,

---

<sup>3</sup> p. 104.

soit « de haut en bas », par *division* du texte total en grands segments, puis de manière récursive de chaque segment en sous-segments.

L'analyse doit-elle être menée en allant « de haut en bas » ou au contraire « de bas en haut »? [...] Les résultats des deux types d'analyse, appliqués aux mêmes matériaux, se recouvrent partiellement, mais, comme de toute façon une procédure unique ne peut jamais suffire, l'alternative de deux procédures cesse d'être cruciale : dans la pratique, les deux se mêlent constamment. Bien entendu, il est très utile d'avoir envisagé de façon précise les conséquences de l'emploi de telle ou telle procédure particulière.<sup>4</sup>

Ainsi, selon Ruwet, *l'organisation intrinsèque de l'œuvre*, puisqu'elle est supposée exister de manière immanente, *prime sur la procédure de découverte de cette organisation*, qui est indépendante de cette organisation. La procédure de découverte ne jouerait donc ici qu'un rôle secondaire qui n'influerait pas directement sur les conditions d'existence des structures découvertes. Notre approche, développée dans ce chapitre, soutient la thèse contraire, suivant laquelle c'est la manière d'appréhender la partition qui détermine la structure découverte, laquelle est *produite* de toute pièce par l'analyse. La manière la plus naturelle et universelle de déterminer cette structure revient alors à appréhender la musique de la même manière que l'écoute, qui procède « de bas en haut », par découverte progressive des structures élémentaires puis de leur inclusion au sein de structures de niveau supérieur. Or la démarche de Ruwet procède, au contraire, principalement « de haut en bas » : des répétitions de longueur maximale sont d'abord déterminés, formant un niveau I ; puis, des répétitions de tailles plus petites forment des niveaux inférieurs II, III, etc. La mise en œuvre d'une telle méthodologie s'avère problématique. D'ailleurs, l'analyse pratique des exemples musicaux destinés initialement à l'illustration de ses principes, par exemple un *Geisslerlied* médiéval<sup>5</sup>, ne semble pas suivre précisément les préceptes formalisés au préalable, mais se fonder au contraire sur des stratégies intuitives non-explicitées. En fait, il peut être montré qu'une mise en application rigoureuse de la méthodologie proposée par Ruwet donne lieu à des résultats curieux qui ne correspondent pas aux intuitions des auditeurs, ce qui invalide le modèle (LARTILLOT 2004c). L'informatique peut, ici aussi, offrir un secours précieux. En effet, une *modélisation* informatique – c'est-à-dire la construction, par l'intermédiaire de programmes informatiques et d'algorithmes, de stratégies de découverte de structures musicales – met explicitement et immédiatement en évidence ses capacités et sa pertinence, et peut être aisément validée – ou invalidée – en fonction de son efficacité.

### **2.3. Comment appréhender les structures musicales.**

Nous proposons de répondre tout à la fois aux exigences de description précise et exhaustive des structures musicales, d'une part, et de corroboration étroite avec les stratégies d'écoute effectives, d'autre part, par l'intermédiaire d'une automatisation informatique de l'analyse musicale suivant une

---

<sup>4</sup> p. 103.

<sup>5</sup> Nous proposerons d'ailleurs une analyse informatique de cette pièce, au chapitre B.2.

modélisation des stratégies d'écoute. Cette modélisation sera élaborée sous la forme d'un ensemble complexe de procédures soigneusement déterminées, lesquelles seront appliquées sur chaque œuvre musicale de manière exhaustive. La grande complexité des opérations mises en jeu, que l'on suppose présente au sein du système cognitif gérant l'écoute musicale, pourra être gérée à l'aide de l'automatisation informatique.

On pourrait alors objecter qu'une modélisation des mécanismes de découverte de structures musicales doit pouvoir être fondée sur des critères objectifs, basés sur la structure même des œuvres musicales, et non sur de simples critères subjectifs fondés sur les possibilités d'appréhension des auditeurs, lesquelles s'avèrent d'ailleurs limitées et peu fiables. L'« analyse du niveau neutre » en particulier, clef de voute méthodologique de la sémiologie musicale de Jean-Jacques NATTIEZ (1975), se penche sur les configurations « immanentes » des œuvres musicales, qui, suivant la tripartition élaborée dans le cadre de cette sémiologie, sont supposées exister indépendamment des niveaux *poïétiques* des intentions du compositeur, et du niveau *esthétique* de la réception de l'œuvre par les auditeurs.

Le niveau neutre est un niveau d'analyse où on ne décide pas *a priori* si les résultats obtenus par une démarche explicite sont pertinents du point de vue de l'esthétique et/ou de la poïétique. Ce qui rend neutre ce niveau descriptif, c'est que les outils utilisés pour le découpage des phénomènes [...] sont exploités systématiquement jusqu'à leurs ultimes conséquences, et ne sont remplacés que lorsque de nouvelles hypothèses ou de nouvelles difficultés conduisent à en proposer de nouveaux. « Neutre » signifie ici que l'on va jusqu'au bout de l'application d'une procédure donnée, indépendamment des résultats obtenus.<sup>6</sup>

Or, dans l'approche de Nattiez, l'analyse du niveau neutre est réalisée sous la forme d'une analyse dite paradigmatique, qui reprend en fait les principes de la méthodologie de Ruwet. Or nous avons montré que cette méthodologie n'est en rien explicite et systématique, mais au contraire fondée en contrebande sur les intuitions du musicologue. En pratique, l'analyse du niveau neutre consiste donc en une segmentation préalable et intuitive du discours musical en unités distinctes, auxquelles seront associées ultérieurement des fonctions. L'absence de critères *a priori* est justifiée par le constat sceptique d'une incapacité de fonder de manière générale une notion de pertinence universelle. Toute spécification *a priori* de la segmentation aurait pour fâcheuse conséquence, selon l'auteur, de restreindre le champ de découverte :

Rappelons seulement que l'écueil du fonctionnalisme linguistique a été de déterminer *a priori* les fonctions, et ainsi, de passer à côté d'un certain nombre de phénomènes qui, pour ne pas répondre à ces fonctions-là, n'en sont pas moins constitutifs de l'objet analysé. Aussi convient-il de ne pas mettre la charrue avant les boeufs et de bien séparer les opérations par lesquelles on identifie les unités et celles par lesquelles on détermine leurs fonctions.<sup>7</sup>

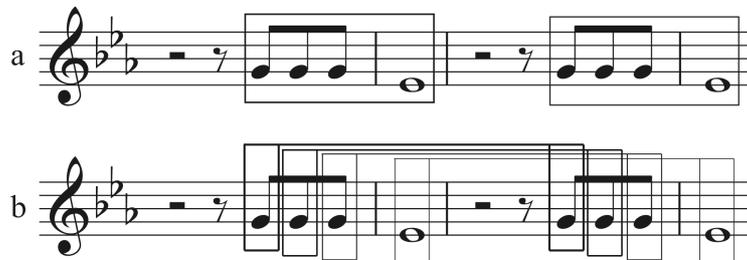
---

<sup>6</sup> p. 54-55.

<sup>7</sup> p. 408.

Là où le fonctionnalisme tente, dans une démarche scientifique, de définir un ensemble de *fonctions constitutives* et de tester leur pertinence par rapport à l'expérience, la méthodologie de Nattiez s'oppose à la modélisation scientifique — qui ne peut, par essence, rendre compte que d'une partie de la réalité — et renonce donc au projet même d'explicitation des critères de découverte.

Il peut être remarqué en fait que la simple définition du concept de motif en tant que séquence de notes contiguës — et non en tant qu'ensemble de notes dispersées çà et là au sein de la partition — provient implicitement de propriétés caractéristiques de l'écoute musicale, fondée sur la mémoire à court terme et sur le mécanisme de rétention (HUSSLERL 1905). Par exemple, une répétition exacte d'un motif (Figure 1a) pourrait être considérée de manière duale sous la forme d'un ensemble de couples de notes (Figure 1b), seconde représentation implicitement rejetée par l'intuition de l'auditeur — et du musicologue — en raison de sa non-prégnance auditive. Ceci soutient l'hypothèse suivant laquelle les structures pertinentes d'une œuvre musicale sont reconstruites suivant une stratégie d'appréhension, et non en vertu de propriétés purement structurelles.



**Figure 1.** L'extrait musical présenté est composé d'un motif de quatre notes répété deux fois de manière exactement identique (ligne a). Ce même extrait peut tout aussi bien être décrit sous la forme de quatre répétitions d'un motif composé de deux mêmes notes distantes de deux mesures (ligne b). Ces deux descriptions sont exactement équivalentes d'un point de vue purement structurel. Cependant, l'intuition musicale, conditionnée par les contraintes cognitives de l'écoute musicale, tend à préférer la description a, et à rejeter la description b, laquelle ne peut pas être perçue en tant que telle.

Le problème se complique une fois posée la question de la découverte de répétitions approximatives. Comme l'explique Emilios CAMBOUROPOULOS (1998) :

Si la similarité (c'est-à-dire la répétition non simplement exacte) doit être prise en compte, alors l'analyse du niveau neutre devient difficilement maniable car une séquence musicale peut être considérée comme similaire de toute autre séquence musicale pour une raison ou une autre. L'analyse du niveau neutre est utile seulement si elle est guidée par une quelconque heuristique — par exemple, fondée sur des principes d'ordre cognitif.<sup>8</sup>

ce qui remet donc en cause la notion même de « neutralité », et soutiendrait au contraire l'hypothèse d'une étroite interdépendance entre structure musicale et stratégies de composition et d'audition.

#### **2.4. Modéliser les stratégies d'écoute**

De prime abord, l'écoute apparaît comme un acte individuel singulier, qui dépend totalement de l'auditeur, de sa volonté, son état d'esprit, son expérience musicale, et des aléas de la psyché. L'écoute offre en effet des représentations variables, pour un même phénomène entendu, et aussi pour

un même auditeur : elle diffère en fonction des compétences développées par l'auditeur et de son degré de concentration. Elle ne semble pas garantir un point de vue fiable et solide sur lequel fonder une quelconque systématisation de l'analyse musicale. Elle présente un caractère « dynamique »<sup>9</sup> : évoluant au cours du temps, elle saisit de manière implicite et peu contrôlée le flux musical en perpétuelle fuite. Elle n'offrirait ainsi qu'une vision linéaire de la musique, une représentation difficilement accessible, car mouvante et floue, qui ne peut être décrite et extériorisée que partiellement. Le fruit de l'écoute n'est apprécié de l'auditeur, pour une grande part, que de manière purement implicite et émotionnelle. Un grand nombre de jugements opérés lors de l'écoute ne font pas l'objet d'une réelle explicitation, et contribuent à l'établissement d'une sensation générale, intellectuelle et émotive, à la fois complexe et floue. Enfin, la totalité de la représentation offerte par l'écoute à l'instant même du phénomène sonore est éphémère : elle ne perdure que pendant un temps limité dans l'esprit de l'auditeur.

L'analyse, au contraire, est une opération formelle et explicite s'appuyant sur une objectivation de l'expression musicale. Elle serait en mesure d'apporter une compréhension globale, libérée de la contrainte temporelle. L'analyse a d'ailleurs pour intérêt principal de guider l'écoute, de lui offrir une vision plus riche et plus explicite, moins fragmentaire et moins fuyante, du flux musical, de mettre en évidence explicitement et clairement ce que l'on peut percevoir de manière implicite et partielle. L'analyse offre ainsi à l'écoute une béquille lui permettant de se développer davantage au sein du champ de ses possibilités.

Si l'acte analytique semble « figé dans le temps », c'est parce que l'objet de l'étude, ainsi que son résultat, sont tout deux représentés généralement dans un espace figé, centré autour de la partition. Malgré tout, l'analyse à proprement parler ne considère pas l'objet musical d'un seul tout : il le parcourt, d'une manière ou d'une autre. Or l'écoute, quant à elle, bénéficie d'une mémoire, qui lui permet elle aussi de parcourir à souhait le phénomène dans sa forme mémorisée, ainsi que la structuration qu'elle constitue sur ce phénomène mémorisé. L'écoute et l'analyse diffèrent alors par les caractéristiques nettement divergentes de leur mémoire respective : l'une reste infidèle, imprécise et aléatoire, l'autre est explicite, objective et figée. Remarquons alors que, d'un autre côté, la mémoire vivante de l'écoute, malgré tous ses défauts, peut néanmoins s'enorgueillir d'une compétence qui a toujours manqué aux mémoires matérielles de l'analyse : elle est *associative*, et peut faire jaillir à tout moment des relations pertinentes d'analogie entre ses éléments. Et c'est justement un tel manque qui pose problème à l'analyse musicale<sup>10</sup>. Mais l'associativité de la mémoire n'est effectivement mise en œuvre que de manière imprécise et aléatoire. En effet, le rappel de cette mémoire offre une certaine

---

<sup>8</sup> Traduction personnelle, p. 9.

<sup>9</sup> Nattiez (1975), p. 73.

<sup>10</sup> L'analyste engage lui aussi sa mémoire associative lorsqu'il tente de déterminer les structures motiviques émergentes. La vision d'ensemble de la partition semble pouvoir aiguïser les capacités de rappel de l'analyste, qui restent cependant fatalement limitées.

variabilité, en particulier en raison des capacités variables d'accessibilité et d'identification. Une modélisation informatique de la mémoire associative permettra alors d'en approfondir les potentiels.

L'écoute est donc considérée, en accord avec les hypothèses cognitivistes (par exemple, DOWLING et HARWOOD 1986, HOWELL, CROSS et WEST 1985), comme un acte cognitif régi par un ensemble de processus formant un système, et présentant des invariances au-delà de son apparente singularité. L'illusion de l'immanence, qui semble avoir aveuglé l'approche structuraliste, ne prouverait-elle pas l'existence de certains automatismes de l'écoute ? Toutes les formes qui semblent immanentes au sein du discours musical sont des structures conçues par le compositeur et perçues par l'auditeur. La trace auditive ou écrite de tels phénomènes ne présente pas de forme en soi, elle n'est qu'un substrat susceptible de provoquer de tels effets. On suppose donc ici que l'écoute n'est pas fondamentalement irrationnelle et singulière, mais semblerait plutôt régie par des mécanismes généraux. L'apparente singularité de l'écoute peut être expliquée par l'importante marge de variabilité de ces mécanismes généraux, notamment les variabilités des capacités de discrimination auditive, de concentration, de mémorisation, mais aussi le savoir et la culture de chaque individu. La singularité de l'écoute individuelle résulte d'un parcours particulier au sein d'un ensemble plus général de compétences potentielles d'écoute. De même, la fluctuation, la non-stabilité et la fugitivité de la représentation offerte par l'écoute peuvent être considérées comme des défauts d'accession à cette représentation potentielle.

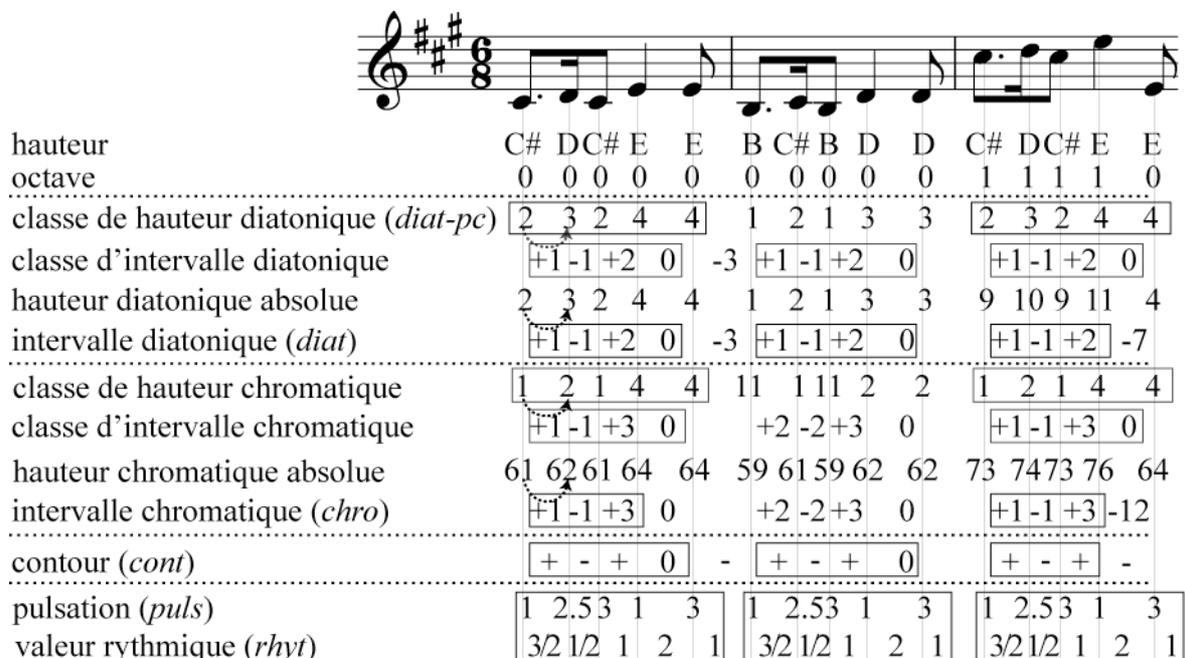
Nous avons montré que l'analyse musicale a pour mission de mettre en évidence des stratégies d'écoute que l'écoute réelle entreprend de manière partielle et principalement implicite. En un mot, l'analyse musicale est tendue vers une transcendance de l'écoute. Si l'analyse musicale est capable d'une telle transcendance, c'est parce qu'elle tire profit des avantages de l'écriture par rapport à l'écoute : notamment sa spatialité, sa stabilité et son explicité. Mais l'écriture ne suffit pas, car elle présente également de nombreux défauts. D'une part, elle ne met pas en évidence les stratégies entreprises par l'écoute. D'autre part, elle ne dispose pas, contrairement à la mémoire humaine, d'un potentiel d'associativité. Les vertus de stabilité et d'explicité de l'écriture peuvent être attendus également d'une écoute envisagée dans sa totalité potentielle. Une systématisation du potentiel de l'écoute assure la transcendance de l'écoute qu'offre également l'écriture, mais apporte également une seconde transcendance par la réalisation potentiellement totale des mécanismes d'associativité. La systématisation de l'analyse suscite parfois de grandes craintes, en raison du caractère machinal et aveugle d'une telle opération, qui peut provoquer un « effacement de l'auditeur » (COOK 1987). Si toutefois cette systématisation consiste, dans l'idéal, en une modélisation de l'ensemble des stratégies d'écoute potentielles, l'auditeur ne se trouve alors plus nié ici, mais intégré au sein d'un vaste ensemble de possibilités.

### 3. Une identification adaptative au sein d'un espace multi-dimensionnel

La réflexion engagée au cours du précédent paragraphe a permis de mettre en évidence la nécessité et l'utilité d'une modélisation automatique des mécanismes cognitifs de découverte de structures musicales. Dans la suite de ce chapitre sont présentés notre proposition de modélisation, ses principes de base et les stratégies sous-jacentes permettant de contrôler certains phénomènes de redondance combinatoire.

#### 3.1. Les dimensions musicales

La complexité de la recherche de répétitions motiviques, dans le cadre musical, résulte en particulier de la multiplicité des dimensions paramétriques suivant lesquelles se construit le discours musical.



**Figure 2.** Description d'une séquence musicale. Les répétitions de suite de valeurs, encadrées, forment des *patterns*.

La hauteur diatonique indique la hauteur de chaque note par rapport à l'échelle tonale ou modale sous-jacente. Les caractéristiques générales de cette échelle (c'est-à-dire la tonalité ou la modalité de la pièce) sont généralement mentionnées explicitement sur la partition. Mais en général, les œuvres musicales présentent des modulations locales non décrites explicitement sur la partition. Ces modulations doivent alors être prises en compte par l'intermédiaire d'une analyse harmonique (tonale ou modale). En particulier, l'analyse d'une partition numérique encodée en format MIDI, laquelle n'inclut aucune description de tonalité,

les hauteurs diatoniques doivent être reconstruites de manière algorithmique (problématique plus connu sous le terme de *pitch spelling*, cf. CAMBOUROPOULOS, 2003; CHEW et CHEN, 2005).

Les *classes de hauteur diatonique* sont représentées sur une échelle de 7 valeurs (de 0 à 6), dont l'origine (0) est associée à la tonique de l'échelle (voir figure 2). La position d'octave est indiquée par un paramètre distinct. La *hauteur diatonique absolue* résulte du produit cartésien des deux dimensions précédentes : chaque valeur de hauteur diatonique absolue est associée à une valeur de classe de hauteur diatonique, à une position d'octave donnée.

Les *classes de hauteur chromatique* sont représentées sur une échelle de 12 valeurs (de 0 à 11), dont l'origine (0) est associée à la hauteur *do*. Ici aussi, l'intégration du paramètre d'octave résulte en une dimension de *hauteur chromatique absolue*. Cette dimension peut être identifiée au paramètre de hauteur utilisé dans le format informatique MIDI suivant lequel la valeur 60 est associée à la hauteur *do* du 4<sup>ème</sup> octave.

Afin de détecter des répétitions de motifs de hauteurs transposés (c'est-à-dire translétés suivant la dimension des hauteurs), les paramètres de hauteurs doivent également être envisagés de manière relative : la position de la hauteur de chaque note est alors considérée en fonction de sa distance (ou *intervalle*) avec la hauteur de la note précédente. Le paramètre d'intervalle peut être défini soit entre hauteurs absolues, ou entre classes de hauteur. On parlera dans le premier cas d'intervalles absolus, et dans le second cas de classes d'intervalles. En particulier, la dimension de classe d'intervalle chromatique est utilisée dans la Pitch-Class-Set theory (FORTE, 1973). Les intervalles de hauteur peuvent également être décrits de manière moins spécifique suivant leur sens de variation, ou *contour* : croissant (la seconde note étant plus haute que la première), décroissant (plus basse) ou monotone (même hauteur). Des études ont montré l'importance perceptive de la dimension de contour (WHITE, 1960; DOWLING et HARWOOD, 1986) : des transformations d'un motif original peuvent être reconnues, même si les valeurs d'intervalles ont été significativement modifiées, tant que le contour reste constant. Cependant, en raison de l'alphabet restreint associé à cette dimension, les répétitions de contour ne peuvent être généralement détectées par les auditeurs si les occurrences sont trop espacées dans le temps (DOWLING et HARWOOD, 1986). L'impact du contour sur l'extraction de motif sera étudié davantage au paragraphe 3.3.

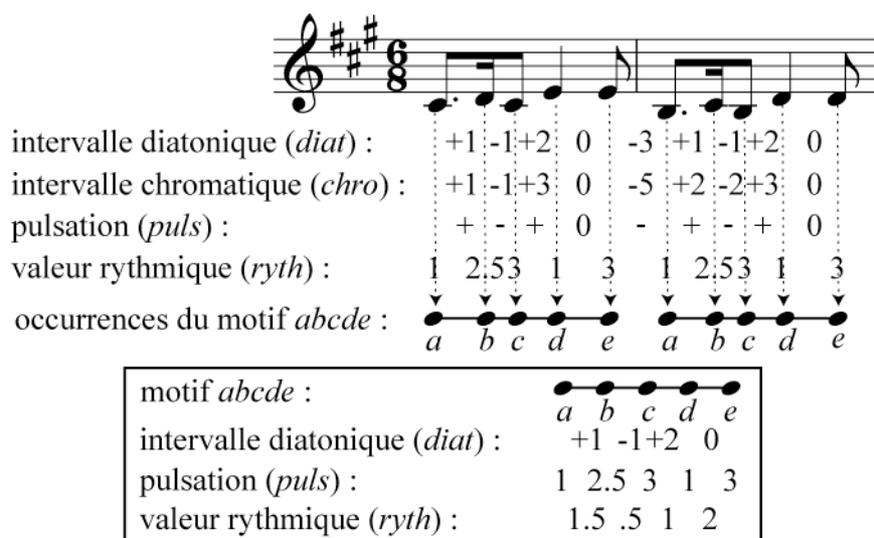
La position métrique indique la phase de chaque note par rapport à la structure métrique. Dans une première approche, cette structure métrique est représentée simplement par une succession de pulsations (temps forts), laquelle est subdivisée en une pulsation deux

fois plus rapide (rythme binaire) ou trois fois plus rapide (rythme ternaire), représentant les temps faibles. Un temps fort est indiqué par la valeur 1, et un temps faible par la valeur 2 pour un rythme binaire, 2 et 3 pour un rythme ternaire. Cette information peut être obtenue directement à partir de la partition. Le format informatique MIDI, par contre, n'explique aucune structuration métrique, qui doit alors être reconstruite par l'intermédiaire d'algorithmes d'inférence de pulsation (TOIVIAINEN, 1997; LARGE et KOLEN, 1994; DANNENBERG et MONT-REYNAUD, 1987), de quantification rythmique (DESAIN et HONING, 1991; CEMGIL et KAPPEN, 2003) et d'induction métrique (TOIVIAINEN et EEROLA, 2006; ECK et CASAGRANDE, 2005). Le résultat de ces algorithmes peut être intégré directement en entrée de notre algorithme d'extraction motivique.

La position métrique joue un rôle important dans le cadre de l'identification rythmique. En particulier, un motif rythmique n'est en général pas détecté si son positionnement sur la pulsation métrique n'est pas conservé. Cette contrainte a été intégrée au sein du modèle.

À chaque note est également associée une *valeur rythmique* (croche, croche pointée, noire, etc.) que l'on notera *ryth*. Cette valeur peut être représentée de manière numérique, sous forme d'un rapport de cette valeur à la pulsation métrique. Dans l'exemple de la figure 2, le dénominateur (8) de la métrique « 6/8 » indiquant que l'unité rythmique est la croche, une noire, valant 2 croches, sera représentée par la valeur rythmique  $ryth = 2$ . Comme une valeur rythmique représente la durée d'une note avant l'apparition de la note suivante ou du silence suivant, nous pouvons la représenter sous la forme d'une distance temporelle entre ces deux événements.

L'ensemble des dimensions musicales prises en compte dans le modèle peut être ordonné de manière hiérarchique. En particulier, le contour est considéré comme plus général que les représentations diatonique et chromatique des intervalles. De nouvelles dimensions musicales peuvent être ajoutées si nécessaire. Leur intégration au sein de l'ensemble nécessite toutefois la définition des relations de dépendances logiques avec les dimensions pré-existantes.



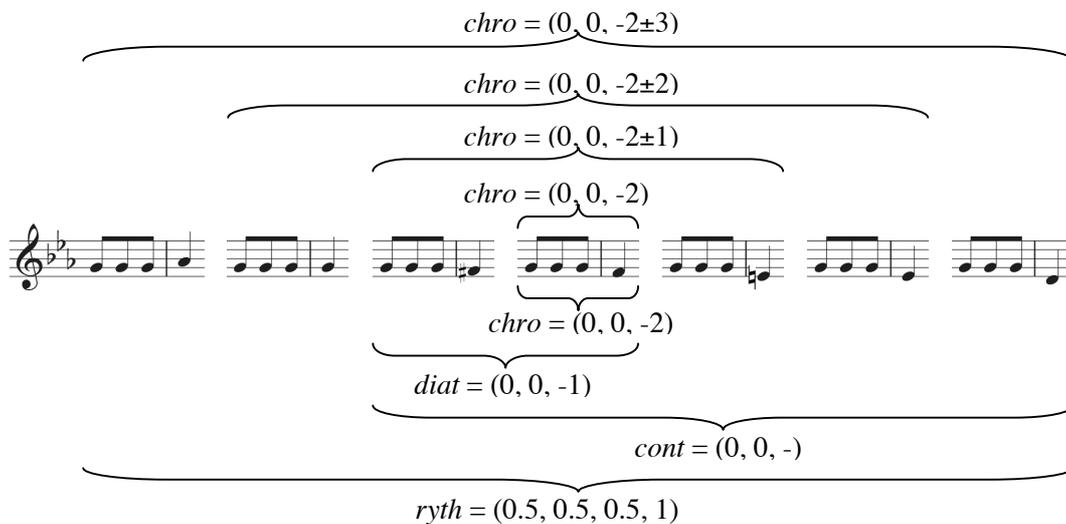
**Figure 3.** Description de la partition suivant différentes dimensions musicales. Les deux motifs encadrés sont deux occurrences d'un motif *abcde* décrit suivant les paramètres d'intervalle diatonique, de pulsation et de valeur rythmique.

Lorsqu'une suite de descriptions est répétée à divers instants d'une pièce musicale, chacune de ces répétitions sera appelée une *occurrence* du *motif* que constitue cette suite de descriptions répétée. Le motif pourra être représenté sous la forme d'une chaîne d'états, et les descriptions du motif seront représentées suivant les transitions successives de cette chaîne (Figure 3). L'ensemble des motifs peut ensuite être représenté sous la forme d'un arbre de préfixes, car deux motifs présentant un même préfixe peuvent être considérés comme deux continuations différentes du même motif que constitue ce préfixe.

### 3.2. L'identification de similarités

Les motifs ne sont généralement pas répétés de manière identique mais subissent différentes formes de transformations locales. Pour pouvoir détecter de tels motifs, il est nécessaire de pouvoir identifier leurs différentes occurrences malgré l'apparente diversité de chacune d'entre elles. Nous avons mis en évidence au sein des approches contemporaines deux stratégies opposées [LARTILLOT 2004a]. Une première méthode, basée sur la notion de *similarité numérique*, consiste à tolérer une certaine marge de dissimilarité entre chacun des paramètres successifs (Cf. par exemple [COPE 1996], [ROLLAND 1999]). Toute la difficulté se concentre alors dans la détermination précise de ce seuil de tolérance, question qui ne semble pas pouvoir trouver de réponse générale. Ainsi, dans l'exemple de la Figure 4 (partie supérieure), les motifs décrits par la succession d'intervalles chromatiques  $chro = (0, 0, -2 \pm 1)$  peuvent être considérés comme similaires au motif décrit par la succession  $chro = (0, 0, -2)$ , car ils présentent une dissimilarité inférieure à un degré. Mais les motifs décrits par la succession  $chro = (0, 0, -2 \pm 2)$  peuvent eux-aussi, dans un moindre degré, être considérés comme similaires au motif  $chro = (0, 0, -2)$ , pour une dissimilarité maximale de deux degrés, et ainsi de suite. Il n'existe donc pas

de véritable dissimilarité maximale, au-delà de laquelle les motifs seraient effectivement *dissimilaires* à  $chro = (0, 0, -2)$ .



**Figure 4.** Deux méthode de classification de motifs : l'une, au-dessus de la partition, basée sur une distance de similarité numérique suivant, par exemple, le paramètre d'intervalle chromatique  $chro$  ; l'autre, en-dessous de la partition, fondée sur une identification exacte le long des différentes dimensions musicales.

Cette première stratégie est remise en cause par certains travaux de référence en cognition musicale [DOWLING et HARWOOD 1986], suivant lesquels la similarité ne doit pas être envisagée en terme de minimisation d'une distance numérique, mais fondée au contraire sur une seconde stratégie d'*identification exacte* le long d'une multiplicité de dimensions musicales de divers niveaux de spécificité. Ainsi, dans le précédent exemple (partie inférieure de la figure 4), le motif  $chro = (0, 0, -2)$  peut être inclus au sein de différents groupements suivant les différentes dimensions musicales, tels que  $diat = (0, 0, -1)$ ,  $cont = (0, 0, -)$  ou  $ryth = (0.5, 0.5, 0.5, 1)$ .

Certaines approches informatiques suivent cette seconde stratégie de détection d'identités le long de différents dimensions musicales (Cf. [CONKLIN et ANAGNOSTOPOULOU 2001], [CAMBOUROPOULOS 2006], [MEREDITH et al 2002]), et recherchent des répétitions le long de chaque dimension ou produit de dimensions. Nous avons alors montré l'existence de motifs qui, parce que construits sur une variabilité de dimensions musicales successives, ne peuvent pas être identifiés par de telles approches [LARTILLOT 2004a]. Par exemple, chaque ligne de la partition de la figure 5 contient une répétition d'un même motif : dans sa première moitié, les paramètres mélodiques et rythmiques sont répétés alors que, dans sa seconde moitié, seuls les paramètres rythmiques sont répétés. La description de ce motif est donc, dans sa première partie, *mélodico-rythmique*, puis, dans sa seconde partie, *rythmique*. Une telle description suivant différents paramètres musicaux successifs sera qualifiée de *description hétérogène*. Notre modèle, contrairement aux approches antérieures, est en mesure de détecter de telles descriptions hétérogènes.

Figure 5 shows two musical staves in G major (one sharp) and 6/8 time. The first staff is a melodic-rhythmic motif, and the second is a purely rhythmic version of the same motif. Below each staff are two rows of data: 'diat' (diatonic intervals) and 'ryth' (rhythmic values). The first five notes of the motif are enclosed in a box in the original image.

diat:	+1	-1	+2	0	-3	+1	-1	+2	0	-3	0	+1	0	+1	+1	-2	
ryth:	1.5	.5	1	2	1	1.5	.5	1	2	1	2	1	2	1	2	1	3

diat:	+1	-1	+2	0	-3	+1	-1	+2	0	-3	+1	+1	+1	-1	-1	-1	
ryth:	1.5	.5	1	2	1	1.5	.5	1	2	1	2	1	2	1	2	1	3

**Figure 5.** Répétition d'un motif hétérogène (encadré) : la première moitié de ce motif est mélodico-rythmique (les intervalles diatoniques et valeurs rythmiques sont répétés) et la seconde moitié simplement rythmique (seules les valeurs rythmiques sont répétées).

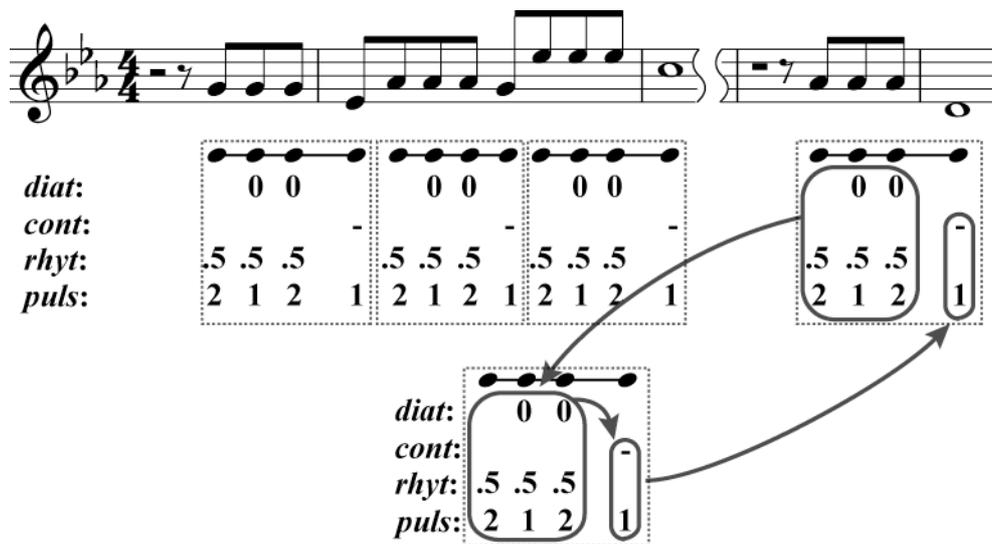
### 3.3. Une solution au problème du contour

Comme mentionné au paragraphe 3.1, en raison du faible degré de spécificité offert par le paramètre de contour, les motifs simplement construits à partir de contours ascendants ou descendants ne sont pas facilement reconnaissables. Il a été suggéré alors que la répétition de séquences de contours ne peut être identifiée que dans un cadre temporel limité : la première occurrence étant toujours conservée dans une mémoire à court-terme lorsque la seconde occurrence est entendue. En effet, les séquences de contours peuvent être plus facilement cherchées au sein d'une mémoire à court-terme qui présente par définition un ensemble assez restreint d'éléments. Une recherche au sein de la mémoire à long terme ne peut en revanche pas être envisagée dans un cadre général, en raison de la taille importante de cette mémoire et de la grande combinatoire de résultats possibles (cf. DOWLING et HARWOOD, 1986). La dimension du contour est d'autant plus restreinte à la mémoire à court-terme que de longs motifs de contours (de 15 notes ou plus par exemple) sont difficilement perçus (EDWORTHY, 1985). Cette restriction entraîne cependant certains paradoxes (DOWLING et FUJITANI, 1971; DOWLING et HARWOOD, 1986). En particulier, si le contour n'a aucun impact sur la mémoire à long terme, comment les différentes occurrences du célèbre motif de quatre notes sont-elles détectées le long du premier mouvement de la *Cinquième Symphonie* de Beethoven (figure 6) ? L'explication suggérée par DOWLING et HARWOOD (1986) est que la répétition multiple du motif permet la mémorisation du contour au sein de la mémoire à long terme. Pourtant, ce motif ne peut-il pas être détecté, en raison de sa construction intrinsèque, même s'il est répété seulement un nombre limité de fois au cours de la pièce (comme figure 6 par exemple) ?

Cette question semble trouver un élément de réponse par l'intermédiaire de la description *hétérogène*, présenté au chapitre précédent, et à travers un découplage du choix des dimensions musicales et de la construction du motif. Une compréhension profonde des propriétés perceptives des motifs nécessite en effet une vision chronologique de la construction de ces structures, en termes d'une concaténation incrémentale des intervalles successifs. La dépendance de cette construction par rapport aux mémoires à court et long terme peut être envisagée sous cet angle. Plus précisément, l'initiation

d'une nouvelle occurrence d'un motif nécessite, comme précédemment, une recherche au sein de la mémoire à long terme. Par contre, on suggère que les extensions ultérieures de la nouvelle occurrence ne nécessitent pas une telle charge computationnelle, car elles ne nécessitent qu'une comparaison entre intervalles nouvellement perçus et les descriptions successives du motif.

Le motif central de quatre notes de la *Cinquième Symphonie* de Beethoven, figure 6, peut être considéré de ce point de vue comme la concaténation de deux intervalles d'unisson (*diat* = 0) suivi d'un contour descendant (*cont* = -). Chaque nouvelle occurrence peut être détectée en raison de la haute spécificité des trois premières notes permettant une recherche au sein de la mémoire à long terme. L'intégration de ces principes au sein du modèle permet une reconstitution du phénomène.



**Figure 6.** Le célèbre motif de quatre notes du début du premier mouvement de la *Cinquième Symphonie* de Beethoven peut être considéré comme une concaténation de deux intervalles d'unisson (*diat* = 0) et d'un contour descendant (*cont* = -), dans un rythme uniforme de croches (*rhyt* = 0.5). Chaque nouvelle occurrence du motif, telle que l'exemple hypothétique présente dans les deux dernières mesures, peut être détectée suivant un processus à deux phases : la description spécifique des deux premiers intervalles entraîne l'appariement avec la mémoire à long terme, alors que la dernière extension du motif peut suivre la description moins spécifique suivant la dimension du contour.

#### 4. Le contrôle de la redondance combinatoire

Une mise en œuvre directe de l'algorithme ainsi présenté sur des exemples musicaux, même simples, donne lieu à la découverte d'un grand nombre de motifs, qui ne correspondent pas, pour la plupart d'entre eux, à des structures effectivement perçues, et conduisent à une explosion combinatoire. Cela est dû d'une part au fait que notre approche accepte un nombre important de configurations différentes, notamment une description hétérogène des motifs. Mais il s'avère que les approches algorithmiques actuelles se heurtent elles aussi à de telles difficultés, et proposent de réduire la complexité par l'intermédiaire d'un filtrage des résultats suivant des critères appliqués globalement, tel qu'une sélection des motifs les plus longs ou les plus fréquents (Cf. [CAMBOUROPOULOS 2006], [CONKLIN et ANAGNOSTOPOULOU 2001], [MEREDITH et al

2002]). Toutefois, un tel filtrage ne permet pas d'améliorer la pertinence perceptive des résultats, mais entraîne en revanche leur appauvrissement.

#### **4.1. Autres approches**

La problématique d'extraction de motifs répétés au sein de séquences musicales symboliques a été envisagée suivant de multiples approches. CAMBOUROPOULOS [2006] procède à une recherche de répétitions exactes au sein de projections suivant diverses dimensions (ou conjonction de dimensions) de l'espace musical. L'ensemble de motifs extraits n'est pas considéré tel quel : au contraire, une segmentation globale de la séquence musicale est estimée par l'intermédiaire d'un moyennage des multiples segmentations élémentaires, aboutissant à une compression avec forte perte de l'information structurelle.

Une autre approche [CONKLIN et ANAGNOSTOPOULOU 2001] consiste en la construction d'un arbre de suffixes suivant une multitude de dimensions musicales distinctes. Ici aussi, en raison de la taille importante de l'ensemble des motifs découverts, un post-traitement sélectionne les motifs présentant un nombre minimal d'occurrences et un degré de pertinence statistique suffisant.

ROLLAND [1999] définit une distance numérique de similarité fondée sur des transformations d'édition entre sous-séquences. Des distances de similarité sont estimées entre tous les couples possibles de sous-séquences, et seules les similarités dépassant un certain seuil prédéfini par l'utilisateur sont sélectionnées. À partir du graphe de similarité résultant, des motifs sont extraits suivant l'algorithme de catégorisation *Star center*. L'ensemble de motifs découverts est réduit encore davantage par l'intermédiaire d'heuristiques de filtrage *a posteriori*. En particulier, seuls les motifs répétés au sein d'un nombre minimum de séquences musicales sont conservés.

Une généralisation de l'extraction motivique à la polyphonie a également été proposée [MEREDITH et al 2002]. Les notes des séquences musicales sont représentées sous forme de points au sein de l'espace bi-dimensionnel hauteur / temps, et les répétitions maximales d'ensemble de points sont extraites. Cette stratégie d'inspiration géométrique ne s'applique toutefois pas aux répétitions mélodiques présentant des variations rythmiques. Des post-traitements ont été ajoutés afin de réduire la redondance combinatoire.

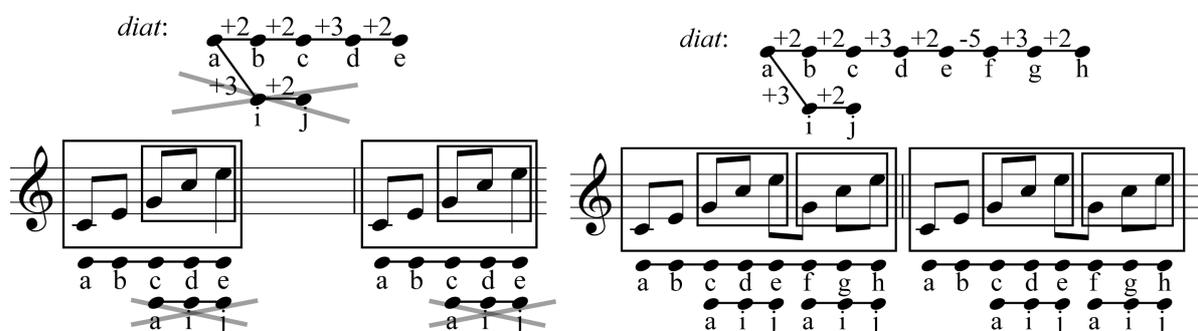
Dans l'ensemble de ces approches, l'explosion combinatoire produite par la phase d'extraction motivique *stricto senso* est filtrée *a posteriori* suivant des critères globaux basés sur la longueur des motifs, leur fréquence, etc. La limitation principale de telles approches est due au manque de sélectivité de ces critères globaux.

Mes travaux permettent alors de mettre en évidence les causes de ce phénomène d'explosion combinatoire et d'inadéquation perceptive des résultats. En fait, la problématique de recherche de répétition conduit implicitement à certaines formes de redondance, que le système cognitif régissant

l'écoute musicale est en mesure de contrôler. Une modélisation informatique de l'écoute musicale se doit alors de reconstruire ces différentes stratégies. Une étude détaillée m'a permis de mettre en évidence une décomposition possible de cette problématique de redondance en un ensemble de sous-problèmes généraux distincts. Sont présentés, dans la suite de cette section, l'ensemble de ces sous-problèmes, et les mécanismes associés conçus dans le but de résoudre chaque sous-problème de la manière la plus simple et générale possible.

#### 4.2. Motifs fermés multi-dimensionnels.

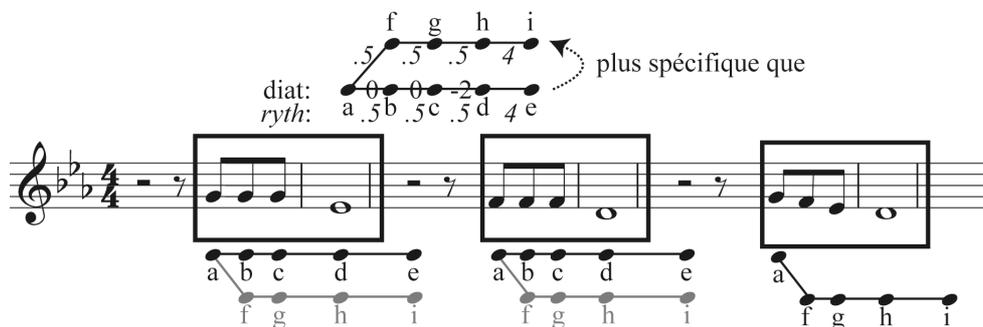
Une première stratégie de réduction combinatoire consiste en une sélection des motifs maximaux [AGRAWAL et SKIRANT 1995 ; ZAKI 2005], c'est-à-dire des motifs non-inclus dans d'autres motifs. Cette heuristique permet une réduction importante de redondance, mais conduit également à une perte importante d'information. En effet, un sous-motif de motif ne doit pas être considéré comme structure purement redondante si son nombre d'occurrences est supérieur à celui du motif qui le contient. Dans ce cas, ces motifs, dits *fermés*, doivent être conservés au sein de la description car ils apportent une information spécifique [ZAKI 2005]. Notre modélisation applique au cadre musical l'heuristique de recherches de motifs fermés. Par exemple, le motif *aij* (Figure 7) est un suffixe du motif *abcde*, qui, dans la partie gauche de la figure, n'existe qu'en tant que simple suffixe, alors qu'il présente, dans la partie droite, des occurrences qui ne sont pas suffixes d'occurrences de *abcde*. Dans ce deuxième cas, le motif *aij* est fermé, et peut désormais être représenté explicitement. Ce principe peut être formalisé de manière plus précise en définissant d'une part, la *classe* d'un motif *m*, comme l'ensemble de ses occurrences, et en introduisant ensuite la notion de *relation d'inclusion* entre classe de motifs. La classe du motif *abcde* et celle de son suffixe *aij* sont considérées, dans l'exemple de gauche, comme *égales* car chaque occurrence de *aij* est un suffixe d'une occurrence de *abcde*. À droite, au contraire, il existe une occurrence de *aij* qui n'est pas un suffixe d'une occurrence de *abcde*. La classe du motif *abcde* est alors dite *strictement incluse* dans celle du motif *aij*. Pour que le suffixe *aij* d'un motif *abcde* puisse être considéré comme un motif autonome, sa classe ne doit alors pas être égale à celle de *abcde*.



**Figure 7.** À gauche, le motif *aij*, de description *diat* = (+3, +2), n'existe qu'en tant que simple suffixe de *abcde* de description *diat* = (+2, +2, +3, +2), car les classes des deux motifs sont égales. Le motif *aij* n'est donc pas représenté explicitement. À droite, en revanche, le motif *aij* n'est pas un simple

suffixe, ni de *abcde*, ni de *abcdefgh*, car leur classe est strictement incluse dans celle de *aij*. Le motif *aij* est donc fermé et peut être conservé.

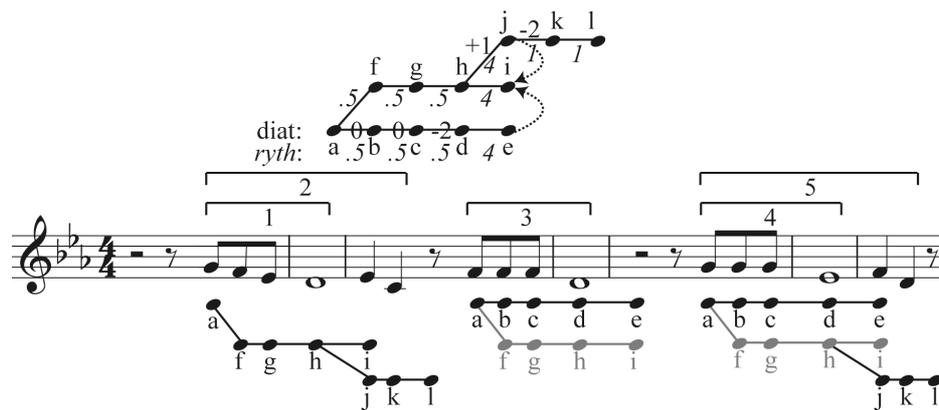
Afin de prendre en compte la multidimensionnalité de l'espace paramétrique musical, il a fallu généraliser la relation d'inclusion entre motifs – qui définit le motif fermé – en l'intégrant dans l'espace paramétrique multidimensionnel caractérisant le domaine musical. Ce problème peut être résolu en considérant la correspondance de Gallois entre les espaces de classes et de descriptions des motifs, étudiée notamment dans les travaux d'Analyse de concepts formels (*Formal Concept Analysis*) [GANTER et WILLE 1999]. Une telle heuristique peut être appliquée à une relation plus générale de *spécificité* entre descriptions. Le motif *abcde* (Figure 8) présente une description mélodico-rythmique  $ryth = (0.5, 0.5, 0.5, 4)$  et  $diat = (0, 0, -2)$ , alors que le motif *afghi* présente la partie strictement rythmique de cette même description  $ryth = (0.5, 0.5, 0.5, 4)$ . Le motif *abcde* peut donc être considéré comme plus spécifique que le motif *afghi*, car sa description contient plus d'information. Le motif moins spécifique *afghi* ne doit généralement pas être explicitement représenté si sa classe est égale à celle du motif plus spécifique *abcde*. Si par contre il existe une occurrence du motif *afghi* qui ne soit pas occurrence du motif *abcde*, comme on peut l'observer dans la partie droite de la figure, alors le motif *afghi* peut être explicitement représenté. En incluant la relation de suffixe au sein de la relation de spécificité – en posant qu'un suffixe d'un motif lui est moins spécifique – alors cette nouvelle heuristique englobe également le filtrage des suffixes décrit dans le précédent paragraphe.



**Figure 8.** Le motif *afghi* de description  $ryth = (0.5, 0.5, 0.5, 4)$  est moins spécifique que le motif *abcde*, qui contient en plus la description  $diat = (0, 0, -2)$ . Au terme des deux premières occurrences, les deux motifs étant de même classe, seule le motif le plus spécifique *abcde* doit être explicitement représenté. Mais une fois découverte la troisième occurrence de *afghi*, qui n'est pas une occurrence de *abcde*, le motif moins spécifique peut lui-aussi être explicitement représenté.

Cet effort d'optimisation de la description motivique a été prolongé, en ajoutant un principe de description maximale spécifique des occurrences de motifs. Lorsque une occurrence d'un motif est découverte (par exemple, le motif *abcde*), toutes les occurrences de motif considérées comme moins spécifiques (par exemple, le motif *afghi*) n'y seront pas superposées, car elles n'apportent aucune information nouvelle, et peuvent être directement déduites à partir de l'occurrence du motif plus spécifique (*abcde*) et de la relation de spécificité (entre *abcde* et *afghi*). Les descriptions moins spécifiques doivent toutefois être prises en compte implicitement car leur extension peut parfois mener à des descriptions spécifiques [LARTILLOT 2005].

Par exemple (Figure 9), les deux groupements 1, 3 et 4 forment des occurrences du motif *afghi*. Les groupements 3 et 4 forment, quant à eux, des occurrences du motif *abcde*. Puisque ce motif est plus spécifique, le motif moins spécifique *afghi* n'a pas besoin d'être explicitement associé au groupement 4 (il est donc représenté en grisé dans la figure). Toutefois, afin de pouvoir détecter que les groupements 2 et 5 forment des occurrences du motif *afghijk*, il est nécessaire de considérer implicitement le motif *afghi* comme occurrence du groupement 4. Cette reconstitution de l'information implicite est effectuée suivant un parcours du réseau des motifs le long de la relation de spécificité.

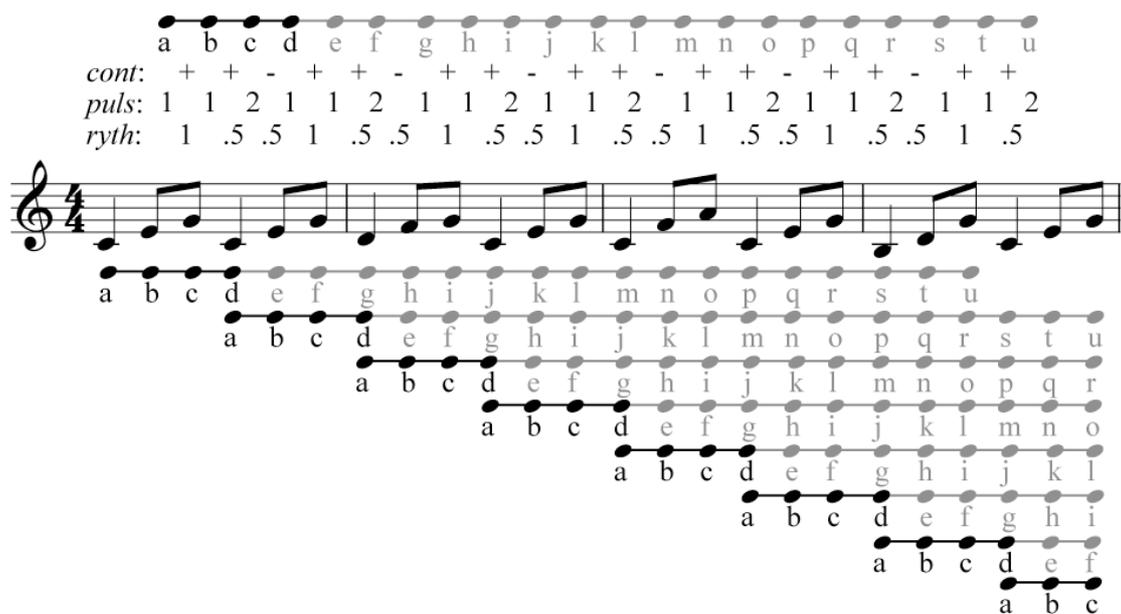


**Figure 9.** Le groupement 4 peut être simplement considéré comme occurrence du motif *abcde*. Toutefois, pour pouvoir découvrir que le groupement 5 est occurrence du motif *afghijk*, il est nécessaire d'inférer implicitement que le groupement 4 est également occurrence du motif *afghi*.

La mise en œuvre de tels principes nécessite une gestion soignée du réseau des motifs, et notamment un parcours de l'ensemble des motifs et de l'ensemble des extensions possibles orienté des plus spécifiques vers les plus généraux. Un tel mécanisme permet ainsi d'assurer une description motivique optimale, par compression sans perte de la taille de la description, suivant un filtrage de la redondance. Une telle optimisation est nécessaire afin, d'une part, d'assurer la clarté des résultats, et d'autre part, de limiter la complexité combinatoire des calculs.

#### 4.3. Une modélisation par cycles des répétitions successives de motifs.

Cette section présente un autre facteur général de redondance combinatoire, maîtrisé par l'intermédiaire d'une approche de type *Gestalt*. Cette explosion combinatoire survient lors de la répétition successive d'un même motif (par exemple le motif rythmique *abcd* à la figure 10). Comme chaque occurrence est poursuivie par le début d'une nouvelle occurrence, chaque motif peut être étendu (aboutissant au motif *abcde*) par un intervalle dont la description est identique à celle du début de ce même motif (c'est-à-dire *ab*). Une telle extension peut alors être prolongée de manière récursive (en *abcdef*, *abcdefg*, etc.), et l'on obtient alors une explosion combinatoire de motifs qui ne sont pas perçus en tant que tels en raison de leur enchevêtrement complexe [CAMBOUROPOULOS 1998].



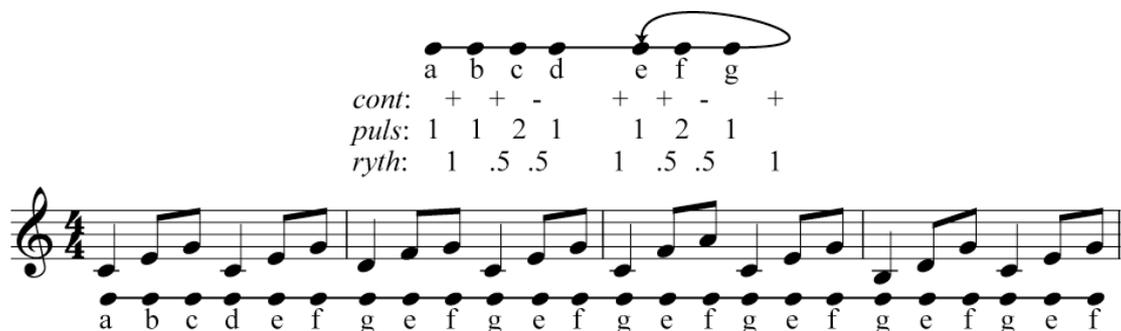
**Figure 10.** La répétition successive multiple du motif *abcd* entraîne en toute logique une extension de ce motif en *abcde*, *abcdef*, etc., formant un enchevêtrement complexe de structures qui ne sont pas perceptibles en tant que telles.

Une telle configuration structurelle est au contraire généralement perçue de manière beaucoup plus simple suivant un principe *Gestalt* de cyclicité motivique. La représentation par graphe de la figure 10 met en effet en évidence le fait que le dernier état de chaque occurrence du motif *abcd* se superpose au premier état (*a*) de l'occurrence suivante. Intuitivement, il semblerait que l'écoute soit encline à fusionner ces deux états, et à opérer un rebouclage du dernier état (*d*) du motif vers le premier (*a*) (Figure 10). Se constitue ainsi un motif cyclique oscillant entre trois états *e*, *f* et *g* à partir du motif acyclique initial *abcd* (Figure 11). Et en effet, lorsque l'on entend la suite de cette séquence, on perçoit effectivement cette progressive oscillation entre ces trois états. Une telle modélisation à l'aide de cycles semble ainsi intuitivement être mise en œuvre lors de l'écoute, et résout de surcroît le problème de redondance combinatoire.

Ce phénomène de répétition successive, bien que très fréquent notamment dans l'expression musicale, a été peu étudié. Une stratégie simple consiste en l'interdiction de tout chevauchement entre motifs [TANAKA et al 2005]. Cette heuristique se fonde cependant sur une contrainte forte – de segmentation stricte et unidimensionnelle de la séquence temporelle – qui n'est généralement pas valable dans le domaine musical. De manière moins contraignante, la minimisation des chevauchements de motifs a été envisagée séparément pour chaque classe de motif [CAMBOUROPOULOS 1998]. Mais une telle sélection, effectuée sur les motifs de manière globale, provoque une excessive suppression de motifs pertinents, alors que notre approche fondée sur les configurations locales entre occurrences de motifs permet un filtrage plus précis. La problématique de recherche automatique de motifs périodiques au sein de séquences temporelles a été étudiée dans des travaux récents [HAN et al. 1998, 1999 ; MA et HELLERSTEIN 2001, YANG et al. 2002]. Ces travaux ne prennent cependant pas en compte l'articulation importante entre motifs périodiques et

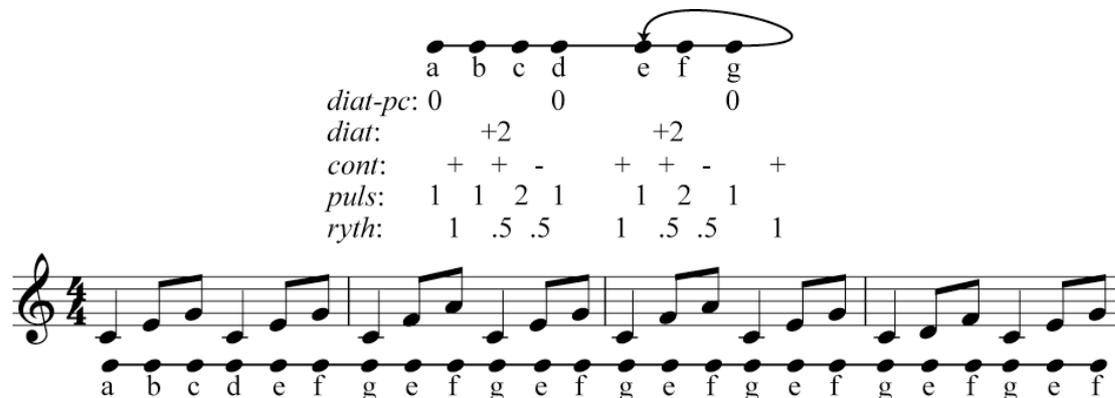
motifs non-périodiques, et ne permettent donc pas de maîtriser le phénomène d'explosion combinatoire.

L'approche proposée ici permet au contraire une articulation étroite entre motifs périodiques et non-périodiques, assurant une réduction sans perte de la redondance combinatoire. Une telle articulation nécessite une lecture strictement *chronologique* de la séquence musicale [LARTILLOT 2005]. La problématique de périodicité a été généralisée à la multidimensionnalité de l'espace paramétrique musical. À cette fin, il a été nécessaire d'étendre la notion de relation de spécificité aux motifs cycliques, et de développer des mécanismes de généralisation et spécification progressives des cycles motiviques, afin de permettre une description la plus compacte possible de la séquence musicale.



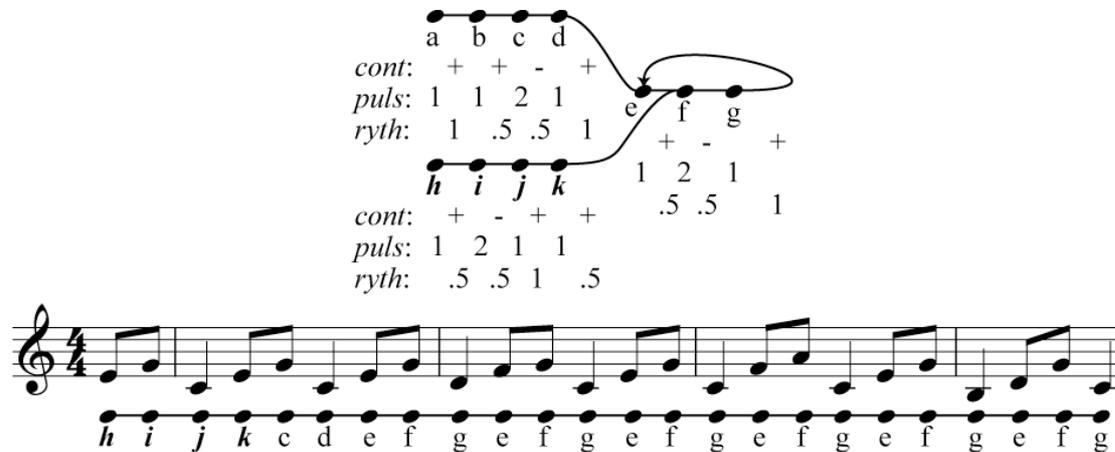
**Figure 11.** L'écoute d'une répétition successive du motif *abcd* induit la perception d'une cyclicité de ce motif, aboutissant à une oscillation entre trois états *e, f* et *g*.

Les motifs cycliques peuvent, eux aussi, être décrits de manière hétérogène, suivant le formalisme présenté au paragraphe 2.2. Par exemple, dans la figure 12, le motif cyclique est moins spécifique que celui présenté à la figure 11 : chaque nouveau cycle commence par une même note *do* (*diat-pc* = 0), et l'intervalle entre les seconde et troisième notes est toujours une tierce majeure ascendante.



**Figure 12.** Un motif cyclique hétérogène, incluant deux descriptions complètes suivant les dimensions rythmiques et de contour, plus deux descriptions locales : la hauteur *do* associée à la première note de chaque période (*diat-pc* = 0), et l'intervalle de hauteur de tierce majeure ascendante entre les seconde et troisième notes de chaque cycle (*diat* = +2).

A également été intégré un mécanisme d'unification de toutes les rotations possibles d'un motif périodique (*efg*, *fge*, *gef*) en un seul motif cyclique. Par exemple, à la figure 13, la séquence périodique présentée à la figure 11 commence sous une phase différente (sur le contre-temps au lieu du temps fort), mais peut toujours être identifiée au même motif cyclique.



**Figure 13.** La séquence périodique débute sur une phase différente (sur un contre-temps au lieu du temps fort), mais peut cependant être identifiée au même motif cyclique (*efg*).

Par définition du motif cyclique, aucune segmentation n'est représentée de manière spécifique entre les répétitions successives. Pourtant, les auditeurs peuvent parfois être enclins à segmenter à une phase spécifique du cycle. Il serait alors intéressant d'estimer les positions de ces segmentations, en considérant les divers facteurs entrant ici en jeu. Une des stratégies suivies par les auditeurs consiste à segmenter suivant la phase du cycle sur laquelle l'occurrence a débuté. Par exemple, à la figure 13, le cycle commençant sur le second temps, chaque segmentation est effectuée sur le second temps des mesures suivantes. Une autre stratégie se fonde sur les discontinuités locales, en privilégiant les segmentations le long des intervalles de hauteurs élevés ou de durées élevées. D'autres heuristiques se fondent sur la construction globale de la pièce et sur la structure métrique.

#### 4.4. La règle de Gestalt figure/fond.

Un autre phénomène de redondance combinatoire survient lorsqu'un motif spécifique se superpose plusieurs fois à un motif cyclique. Un motif spécifique (par exemple, le motif *ab*, Figure 14, associé à la description mélodico-rythmique  $ryth = 1$  et  $diat = -1$ ), se superpose plusieurs fois sur un motif cyclique (par exemple, le motif simplement rythmique  $ryth = 1$  en bas de la figure 14). En toute logique, le motif spécifique peut être étendu suivant les extensions successives du motif cyclique (aboutissant aux motifs *abc*, *abcd*, etc.). Un tel phénomène peut être rencontré fréquemment au sein de la partition, ce qui entraîne une prolifération combinatoire de structures redondantes. Or il s'avère qu'une telle extension n'est pas réellement perçue par l'auditeur en raison de la règle de Gestalt figure/fond : le motif spécifique (*ab*) forme une *figure*, au-dessus du *fond* que constitue le motif cyclique. Suivant cette règle, la figure ne peut être prolongée (en *abc*) par une description  $ryth = 1$  qui

s'identifie au simple prolongement du fond. Cette règle permet ainsi de gérer de manière sélective les extensions redondantes.

**Figure 14.** Le motif *ab* est une *figure* spécifique au-dessus du *fond* constitué par le motif cyclique, représenté en-dessous de la partition. Suivant la règle de *Gestalt* figure/fond, la figure ne peut être prolongée par une extension (*abc*) qui s'identifie au fond.

#### 4.4. Un système complexe

Le modèle global se décompose en un ensemble de modules dédiés aux différentes problématiques, lesquels sont eux-mêmes décomposés en un ensemble de sous-modules centrés sur des sous-problèmes distincts, et ainsi de suite. Chacun de ces modules et sous-modules peuvent être remodelés et articulés les uns les autres d'une manière flexible, ce qui permet de tester diverses hypothèse de modélisation. La représentation des données a également été conçue dans l'objectif d'offrir un maximum de flexibilité concernant le choix de représentations des structures. Le principe général de la méthodologie de notre projet consiste en une construction progressive du système computationnel. A chaque étape, le comportement général du système est contrôlé, ses éventuels défauts sont diagnostiqués par l'intermédiaire d'une analyse approfondie du comportement de chaque module. L'amélioration de la modélisation s'effectue alors par la correction d'éventuelles erreurs de modélisation, le réglage de certains paramètres, ou l'inférence de nouvelles heuristiques générales.

Les mécanismes de filtrage de la redondance assurent une description motivique optimale. L'information est compressée sans perte, les structures filtrées pouvant être reconstruites de manière implicite. Le filtrage des structures redondantes améliore la clarté des résultats, et en même temps décroît la complexité combinatoire des calculs.

La complexité algorithmique s'exprime sous deux formes. D'une part, par la complexité des structures découvertes : la prolifération de motifs redondants, par exemple, induit une explosion combinatoire, puisque chaque nouvelle structure nécessite des processus permettant d'estimer leur impact sur les autres structures et d'envisager leur éventuelle extension. Assurer une description la plus compacte possible des structures permet ainsi tout à la fois d'offrir des résultats clairs et proches des motifs perçus, et de réduire les risques d'explosion combinatoire. La complexité algorithmique s'exprime d'autre part à travers la mise en œuvre technique de la modélisation. De ce point de vue, la réalisation proposée se présente sous la forme d'un prototype qui demande à être davantage optimisé. La modélisation a toutefois été conçue dans un souci constant de minimisation des coûts

computationnels. Ainsi la détection de répétitions d'intervalles s'effectue par simple identification au sein de listes inversées, ce qui permet de simplifier la procédure élémentaire de recherche de motifs (cf. paragraphe C.1).

#### 4.5. Résultats de la première phase de la modélisation

Cette modélisation a été initialement développée sous la forme d'un système d'analyse motivique automatisée écrit en *Common Lisp* et intitulé *OMkanthus*, intégrée au sein du logiciel *OpenMusic* de programmation graphique de structures musicales [ASSAYAG et al 1999]. Une autre version est intégrée au sein de la panoplie d'outils informatiques dédiés à l'analyse musicale, intitulée *MIDItoolbox* [EEROLA et TOIVIAINEN 2004], conçue en tant qu'extension du logiciel de calculs mathématiques *Matlab*. Notre modélisation est en mesure d'analyser des pièces musicales monodiques – c'est-à-dire constituées d'une suite de notes temporellement non superposées – et de mettre en évidence sous forme graphique les structures découvertes sur une partition.

Le modèle a été testé sur un ensemble d'œuvres musicales de styles variés. Quelques résultats sont présentés dans le tableau 1. Les résultats ont été comparés avec une analyse manuelle réalisée par un musicologue. Des exemples d'analyse sont présentés en annexe B. Une procédure de test est en projet d'élaboration sur une base de donnée musicale de grande taille, afin d'établir une mesure quantitative précise du comportement algorithmique.

TABLE 1 – Résultats d'analyses effectuées par *kanthus* version 0.6.8.

Nom du morceau	Taille du morceau	Classes de motifs découvertes	Classes de motifs pertinentes	Temps de calcul
<i>Geisslerlied</i> (chanson médiévale)	108 notes	6	5 (83%)	2.2 secondes
Mozart, <i>Sonate K 331</i> 1 <sup>er</sup> mouvement, 1 <sup>er</sup> thème	36 notes	14	10 (71%)	0.8 seconde
Bach, <i>Invention BWV</i> 775	283 notes	49	34 (69 %)	37.6 secondes
<i>Au clair de la lune</i>	44 notes	26	20 (77%)	5.6 secondes
The Beatles, <i>Obla Di, Obla Da</i>	390 notes	14	10 (71%)	28.1 secondes

Les approches informatiques actuelles ne sont pas en mesure de présenter un tel degré de précision, ni d'assurer la pertinence perceptive de leurs découvertes. Les analyses produites par le modèle proposé restent cependant encore assez limitées, car de nombreuses composantes de l'expression musicale n'ont pas encore été prises en compte.

## 5. Étude de l'improvisation modale arabe

L'application à des styles musicaux de cultures très différentes (telles que la musique classique occidentale, la musique improvisée arabe, le jazz, ou la musique traditionnelle finnoise) questionne l'emprise culturelle de notre formalisation, c'est-à-dire de vérifier si le domaine de validité du modèle se cantonne à la musique occidentale, ou s'il présente certaines généralités d'ordre cognitive le rendant à même d'analyser des musiques de diverses cultures. Ceci permet de mettre en évidence ce qui relève, dans cette modélisation, du culturel, et ce qui relève au contraire de l'universel.

L'étude d'improvisations modales arabes se restreint à un sous-ensemble de l'espace paramétrique musical défini en paragraphe 3.1. Les *hauteurs diatoniques* sont effectivement prises en compte. L'étude s'est focalisée pour l'instant sur la première partie de l'improvisation *Istikhbar Mhayyer Sika*, construite essentiellement sur un seul genre modal (*Tba' Mhayyer Sika ré*) et ne comporte donc pas de modulation modale. L'estimation de la hauteur diatonique s'effectue donc par l'intermédiaire d'une simple mise en correspondance de la hauteur de chaque note avec la position correspondante sur le mode *Tba' Mhayyer Sika ré* : à ré est associée la valeur d'origine 0, à mi la valeur 1, fa 2, etc. Les *intervalles de hauteurs diatoniques* sont également pris en compte, afin de mettre en évidence la présence de motifs transposés. Les *hauteurs chromatiques* sont intégrées mais n'offrent pas d'intérêt particulier ici, de même pour les *intervalles de hauteurs chromatiques*. Le *contour* est pris en compte également.

Les paramètres rythmiques (*pulsation et valeur rythmique*), par contre, sont exclus de notre étude. En effet, ces improvisations se caractérisent par l'absence de structure métrique. Il n'est donc pas possible de traduire les durées des notes sous forme de proportions rythmiques précises. La transcription proposée par Mondher Ayari offre une description rythmique hypothétique de cette improvisation. Dans notre modélisation, nous avons tenté de quantifier les différentes durées rythmiques en un ensemble simple de durées. Cette solution provisoire ne permet pas d'identifier toutes les répétitions rythmiques. Nous tenterons, dans la suite de nos recherches, de modéliser plus finement ce mécanisme d'identification en autorisant des variations rythmiques.

### 5.1. Modéliser les transformations musicales.

Un motif, formé d'une suite de descriptions le long de divers paramètres musicaux, est découvert lorsque une telle suite est explicitement présente à divers instants d'une séquence musicale. Or une caractéristique essentielle du discours musical est l'existence de mécanismes de transformation musicale, suivant lesquels une suite de descriptions, sous une forme originare simple, se modifie en une suite de descriptions plus complexe. Certaines de ces transformations sont en quelque sorte *réversibles*. C'est-à-dire que l'auditeur peut effectuer l'opération inverse : à partir du niveau de surface que constitue la suite de descriptions ornementée, il peut retrouver la structure profonde que constitue la suite de descriptions originare. Il est nécessaire d'incorporer de telles classes de transformations

musicales au sein de la modélisation, afin de pouvoir rendre possible la détection de motifs exprimés au niveau de la structure profonde, à partir d'une description de l'œuvre au niveau de la surface.

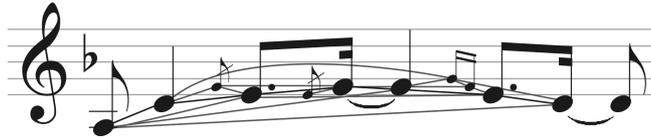
En particulier, un motif peut être *ornementé* par l'ajout de notes supplémentaires dans un voisinage proche (cf. figure 15). Ces ornements n'appartiennent pas à la description du motif, mais ont pour fonction de mettre en valeur les notes auprès desquelles elles sont situées.



**FIGURE 15** – Répétition d'un motif avec ornementation : la deuxième occurrence du motif est enrichie de notes supplémentaires qu'il est nécessaire de filtrer afin de retrouver la répétition de la structure sous-jacente.

Des solutions à cette problématique de recherche de motifs de notes non immédiatement successives ont été proposées (Cf. [MONGEAU et SANKOFF 1990], [ROLLAND 1999], [DANNENBERG et HU 2002]). Elles se basent principalement sur des techniques de programmation dynamique, qui permettent un alignement optimal entre notes similaires de chaque occurrence et se fonde sur une distance d'édition entre séquences musicales. Or, comme expliqué au paragraphe 3.2, l'utilisation d'une distance numérique introduit des difficultés supplémentaires. Cela conduit notamment à un assujettissement de la décision d'identification à un seuil de similarité numérique, qui ne peut être défini que de manière subjective et remet donc en cause la robustesse de l'analyse.

Une nouvelle méthode de résolution est actuellement en cours d'élaboration, fondée sur une identification adaptative, et sur une réduction de l'espace de solutions par l'intermédiaire d'heuristiques basées sur des considérations d'ordre perceptuel. Pour cela, sur la *chaîne syntagmatique* – que constitue la succession apparente de notes de la séquence musicale – est construit un *graphe syntagmatique* contenant l'ensemble de toutes les connexions perceptibles entre notes temporellement proches (cf. figure 16). Des hypothèses d'ordre cognitif et des expérimentations informatiques permettent de mettre en place un ensemble de règles limitant la combinatoire du graphe syntagmatique. L'algorithme d'extraction motivique, initialement conçu pour l'analyse de chaînes syntagmatiques, est généralisé à ce nouveau cadre, à travers un parcours de toutes les chaînes contenues au sein du graphe syntagmatique. Les heuristiques de filtrage adaptatif de la redondance combinatoire, présenté en section 4, ont été généralisées à ce cadre plus large par l'adjonction de nouveaux opérateurs.



**FIGURE 16** – Afin de détecter la répétition de la structure profonde, un graphe syntagmatique (indiqué en gris) est construit, et la recherche de répétition motivique est effectuée le long des différentes branches

### 5.1.1. Construction du Graphe Syntagmatique

Une première stratégie naïve consisterait à construire toutes les relations syntagmatiques possibles entre couples de notes de la partition, afin d’assurer l’exhaustivité des résultats. Mais une telle solution conduit de manière évidente à une explosion combinatoire et n’assure pas la pertinence perceptive des résultats. Afin d’améliorer ces résultats, une série d’heuristiques générales ont été progressivement intégrées au sein du modèle, relatives à divers aspects de la cognition musicale. Dans l’état actuel de nos recherches, encore dans une phase exploratoire, cinq heuristiques générales ont été mises en évidence.

#### *Mémoire à court-terme*

Les relations syntagmatiques sont perçues entre notes qui appartiennent à un même voisinage temporel. Chaque nouvelle note entendue est connectée aux notes précédemment détectées restant encore actives au sein de la mémoire à court terme. Dans la version actuelle du modèle, cette limitation est définie par l’intermédiaire d’un intervalle de temps. Un affinage de ces caractéristiques nécessitera un approfondissement de cette étude.

#### *Notes de passage*

Lorsqu’une note  $n_i$  est entourée de deux notes  $n_{i-1}$  et  $n_{i+1}$  telles que :

- elles forment un même intervalle diatonique,

$$diat(n_{i-1}, n_i) = diat(n_i, n_{i+1}),$$

- l’écart temporel entre  $n_i$  et  $n_{i+1}$  est assez similaire à l’écart temporel entre  $n_{i-1}$  et  $n_i$  :

$$abs(\log(rhyt(n_i, n_{i+1})/rhyt(n_{i-1}, n_i))) < \partial,$$

(où  $\partial$  est fixé à 0.1 dans la modélisation actuelle), alors la note intermédiaire  $n_i$  est considérée comme *note de passage* entre  $n_{i-1}$  et  $n_{i+1}$  : elle est perçue comme un état intermédiaire de l’intervalle entre les deux notes extrêmes, et n’est pas prise en compte en dehors de cette configuration locale. En conséquence, la note de passage  $n_i$  n’est pas connectée syntagmatiquement avec les notes avant  $n_{i-1}$  et après  $n_{i+1}$ .

Par exemple, à la figure 17, la note 7 est une note de passage entre les notes 6 et 8, et n'est donc pas connectée aux autres notes du voisinage, telles que les notes 4 ou 10, bien qu'elles soient de même hauteur (*sol*).

#### *Méta-notes*

Des notes de passage de même hauteur établissent entre-elles une configuration particulière qui semble provoquer un effet de masquage au sein du graphe syntagmatique. En effet, en raison de leur identité paramétrique, elles peuvent être considérées comme des parties successives d'une même méta-note globale. Une des notes est choisie comme représentant principal de cette méta-note. Il s'agit généralement de la première note de la série, ou de la note la plus longue ou la plus accentuée. La position temporelle de cette méta-note est identifiée à la position temporelle du représentant principal. Seule la méta-note, et non pas les notes qui la composent, est connectée avec les autres notes antérieures et postérieures à cette série.

#### *Attraction mélodique*

Le masquage dû aux relations d'unisson, présenté au paragraphe précédent, semble pouvoir être généralisé aux intervalles de demi-tons : deux notes  $n_i$  et  $n_j$  ne sont pas connectées s'il existe une note intermédiaire  $n_k$  formant un intervalle de demi-tons avec la note antérieure  $n_i$ , et de plus longue durée. Par exemple, à la figure, les notes 9 et 13 ne sont pas connectées en raison de la résolution de la note 9 sur la note 11, un demi-ton plus bas.

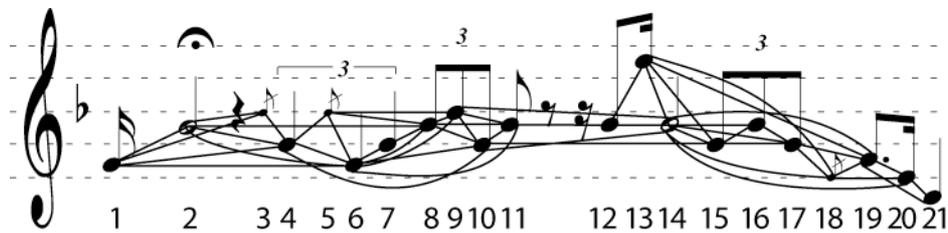
#### *Extrema locaux et progressions internes*

Les notes situées à des extrema locaux du profil de hauteur peuvent plus facilement former des relations syntagmatiques. Dans les autres cas, les relations syntagmatiques seront inférées seulement si les notes intermédiaires ne *disturbent* pas le *domaine de hauteurs* défini par les notes extrêmes. Plus précisément, si la hauteur de la note intermédiaire se situe entre les hauteurs des deux notes extrêmes, et si sa durée est plus longue que la durée de la note finale, alors la note intermédiaire masque la relation syntagmatique entre les deux notes extrêmes.

#### *Exemple*

Ces cinq règles de filtrage permettent une simplification du graphe syntagmatique, lequel met en évidence des relations de successivité entre notes non nécessairement immédiatement successives au sein de la monodie. Figure 17 montre un exemple de construction de graphe syntagmatique à partir d'un extrait de l'improvisation *Istikhbar Mhayer Sika*. Comme on peut le voir, le graphe met en évidence les relations syntagmatiques saillantes entre notes locales. Une validation objective sera mise en œuvre une fois la conception de l'algorithme achevée. Cette étude reste en effet pour l'instant

dans un stade exploratoire. D'autres heuristiques doivent être intégrées, notamment celles fondées sur la notion de pulsation.



**Figure 17.** Construction du graphe syntagmatique sur le début de l'improvisation *Istikhbar Mhayyer Sika*. L'axe vertical est dilaté afin d'améliorer la lisibilité du graphe.

### 5.1.2. Découverte de motifs fermés au sein de graphes syntagmatiques

La procédure de recherche de motifs est mise en œuvre directement sur le graphe syntagmatique. Le graphe est parcouru dans un ordre chronologique, conformément à la description de la section C.1. Plus précisément, à chaque instant  $t$ , la nouvelle note  $n_t$  jouée est ajoutée au graphe et les relations syntagmatiques entre cette note et la note précédente  $n_{t-1}$ , et éventuellement des notes plus anciennes  $n_{t-k}$ ,  $k \geq 1$ , sont ajoutées au graphe. Chacune de ces précédentes notes  $n_{t-k}$ ,  $k \geq 1$  terminent un certain nombre de motifs  $P$ , qui peuvent maintenant être éventuellement étendus par la nouvelle note  $n_t$ . Pour cela, l'intervalle entre  $n_{t-k}$  et  $n_t$  est mémorisé au sein de la mémoire associative du motif  $P$ . Cet intervalle peut soit être reconnue comme une extension déjà connue du motif  $P$ . Sinon, une nouvelle extension est découverte si un autre intervalle similaire existe déjà dans la mémoire associative d'extension du motif  $P$ .

#### *Généralisation de la relation d'inclusion*

Le filtrage adaptatif de la redondance combinatoire est assuré, comme expliquée en section 4, par l'intermédiaire d'une sélection des motifs fermés. Cette heuristique est fondée sur une *relation d'inclusion* entre motifs. Notre étude se limitait à l'analyse de monodies strictes, formées par une simple chaîne syntagmatique, la relation d'inclusion pouvait alors se limiter aux relations de préfixe et de suffixe. Dans le cadre plus général des graphes syntagmatiques, la relation d'inclusion ne peut plus être identifiée simplement aux relations de préfixe et de suffixe : les insertions et suppressions de notes doivent être prises en compte également. Plus précisément, un motif  $P$  est inclus dans un autre motif  $Q$  si les successions de descriptions correspondantes peuvent être superposées, et si la description du motif  $P$  peut être déduite directement à partir de la description du motif  $Q$ .

Pour cela, il est nécessaire de distinguer parmi les différentes descriptions associées à un état donné d'un motif  $P$ , les *descriptions dites de notes*  $N(P)$  qui sont associées à chaque note du motif, et les *descriptions dites d'intervalles*  $I(P)$  associées aux intervalles entre notes successives du motif. Au sein de l'espace paramétrique présenté au chapitre 3.1, les descriptions de notes sont la hauteur des

notes et la pulsation, et les descriptions d'intervalles sont les intervalles de hauteur, le contour et la valeur rythmique.

Le motif  $P$  est inclus dans le motif  $Q$ , ce qui sera noté  $P < Q$ , s'il existe une correspondance strictement croissante  $f$  entre les indices d'états de  $P$  et certains indices d'états de  $Q$  telle que :

- les descriptions de notes de  $P$  sont incluses au sein des descriptions de notes de  $Q$  :

$$\text{pour tout } i \text{ dans } [1, m], N(P_i) \text{ inclus dans } N(Q_{f(i)})$$

- les descriptions d'intervalles de  $P$  sont incluses au sein des cumulations des descriptions d'intervalles correspondants de  $Q$  :

$$\text{pour tout } i \text{ dans } [1, m-1], I(P_i, P_{i+1}) \text{ inclus dans } I(Q_{f(i)}, Q_{f(i+1)})$$

où :

$$I(Q_{f(i)}, Q_{f(i+1)}) = \text{Somme, pour } j=f(i), \dots, f(i+1)-1, \text{ de : } I(Q_j, Q_{j+1})$$

Chaque intervalle est décrit le long d'un ensemble de dimensions musicales. En raison de la propriété d'hétérogénéité définie à la section 3.2, cet ensemble de dimensions musicales n'est pas nécessairement le même pour les différents intervalles successifs constituant un motif donné. La description de la cumulation totale  $I(Q_{f(i)}, Q_{f(i+1)})$  ne contient que les dimensions musicales définies sur la totalité des intervalles constituant la cumulation. Par exemple, si l'un de ces intervalles ne contient aucune description mélodique, alors aucune description mélodique n'est associée à l'intervalle mélodique de la cumulation totale. Dans ce cas, le motif  $P$  sera inclus dans le motif  $Q$  si l'intervalle mélodique de  $Q$  ne contient pas de description mélodique non plus.

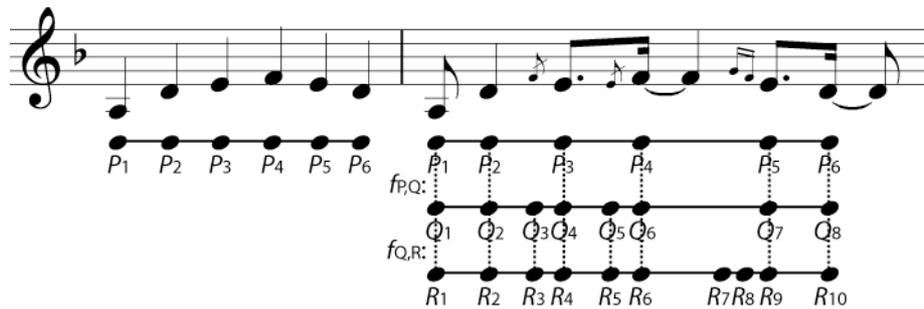
La correspondance  $f$  joue un rôle important au sein du modèle car, comme nous allons le voir, elle permet une reconstruction directe des descriptions générales à partir des descriptions spécifiques. Cette fonction sera appelée *schéma de généralisation* entre les motifs  $P$  et  $Q$ , et notée  $f_{P,Q}$ .

### **Construction du graphe de spécificité**

Comme expliqué précédemment, le modèle inclut un mécanisme de description optimale de la partition, qui filtre de manière adaptative toutes les descriptions qui, bien que fermées, sont redondantes avec certaines descriptions plus spécifiques. Ceci permet une restriction des calculs aux structures les plus spécifiques. Les descriptions plus générales peuvent être reconstruites si nécessaire par l'intermédiaire d'un parcours du graphe de spécificité. Afin de prendre en compte le problème des ornements mélodiques, ainsi que de la polyphonie de manière plus générale, les mécanismes de construction du graphe de spécificité doivent être révisés.

*Transitivité de la relation d'inclusion.*

Le graphe de spécificité est optimisé par l'intermédiaire d'un élagage des relations redondantes. Le calcul des motifs les plus spécifiques et les plus généraux nécessite le parcours d'un nombre multiple de transitions successives au sein du graphe de spécificité. Il est nécessaire ici d'évaluer les conséquences de la propriété de transitivité de la relation d'inclusion sur les schémas de généralisation. Soit  $P$ ,  $Q$  et  $R$  trois motifs de niveaux croissants de spécificité tels que  $P < Q < R$ . Le schéma de généralisation  $f_{P,R}$  est la composition des schémas de généralisation  $f_{P,Q}$  et  $f_{Q,R}$ :  $f_{P,R} = f_{P,Q} \circ f_{Q,R}$ .



**Figure 18.** Relation de spécificité entre un motif ornementé  $R$  et un motif non-ornementé  $P$ , articulée à l'aide d'un schéma de généralisation.

*Mise à jour du graphe de spécificité.*

La relation d'inclusion généralisée définie au paragraphe précédent doit être intégrée au sein du modèle. Suivant notre approche incrémentale, les motifs sont construits suivant un parcours chronologique de la partition. Pour chaque note, les motifs qu'elle termine doivent être construits suivant un ordre décroissant de spécificité. Pour cela, les préfixes candidats à une extension sont considérés suivant un ordre décroissant de spécificité. Cet ensemble ordonné de préfixes-candidats est obtenu en considérant toutes les occurrences de motifs terminées par la note précédente  $n_{i-1}$ .

Chaque fois qu'un nouvel état est ajouté à l'arbre de motif, un nouveau réseau de relations d'inclusions est construit, lequel met en évidence les interdépendances entre les nouveaux états et les anciens états de l'arbre. Le réseau de spécificité est construit de manière incrémentale et récursive : pour chaque nouvel état ajouté au sein de l'arbre de motifs, le réseau est mis à jour à l'aide du réseau associé au préfixe précédent, et de la description de la nouvelle extension.

Soit  $P_i$  le nouveau motif. À son précédent préfixe  $P_{i-1}$  est déjà associé un ensemble de motifs plus généraux  $G(P_{i-1})$  ainsi qu'un ensemble de motifs plus spécifiques  $S(P_{i-1})$ . Les motifs plus généraux que le nouveau motif  $P_i$ , forment le nouvel ensemble  $G(P_i)$ , composé des motifs  $G_{j-1}$  de  $G(P_{i-1})$  qui sont prolongés par une description plus générale que la dernière description de  $P_i$ , c'est-à-dire :

$$G(P_i) = \{G_j \mid G_{j-1} \text{ appartient à } G(P_{i-1}) \text{ et } G_j \text{ inclus dans } P_i\}$$

La construction de l'ensemble des motifs plus spécifiques que  $P_i$ , ou  $S(P_i)$ , suit une description exactement symétrique :

$$S(P_i) = \{S_j \mid S_{j-1} \text{ appartient à } S(P_{i-1}) \text{ et } P_i \text{ inclus dans } S_j\}$$

L'application aux graphes syntagmatiques de cette construction incrémentale des graphes de spécificité nécessite quelques améliorations, qui cassent la parfaite symétrie des opérations incrémentales. En effet, l'ensemble des motifs plus généraux  $G(P_i)$  est composé des extensions directes, non seulement des motifs  $G_{j-1}$  de  $G(P_{i-1})$  plus généraux que le dernier préfixe de  $P_i$ , mais aussi plus généralement des motifs  $G_{j-1}$  de  $G(P_{i-k})$ ,  $k \geq 1$  plus généraux que *tout* préfixe de  $P_i$  :

$$G(P_i) = \{G_j \mid G_{j-1} \text{ appartient à } G(P_{i-k}), k \geq 1, \text{ et } G_j \text{ inclus dans } P_i\}$$

Au contraire, l'ensemble des motifs plus spécifiques  $S(P_i)$  est composé des extensions *non nécessairement directes*  $S_{j+l}$ ,  $l > 0$  des motifs  $S_{j-1}$  de  $S(P_{i-1})$  plus spécifique que le *dernier* préfixe de  $P_i$  :

$$S(P_i) = \{S_{j+l}, l > 0 \mid S_{j-1} \text{ appartient à } S(P_{i-1}) \text{ et } P_i \text{ inclus dans } S_{j+l}\}$$

### 5.1.3. Application aux motifs cycliques.

Cette nouvelle définition de la relation de spécificité peut également être appliquée directement aux graphes cycliques. De nouvelles configurations intéressantes peuvent être ainsi mises en évidence. Par exemple, une pulsation de croches peut être considérée comme plus spécifique qu'une pulsation de noires. Lorsqu'une pulsation de croches est entendue, la pulsation de noires moins spécifique n'a pas besoin d'être représentée puisqu'elle peut être directement reconstruite par l'intermédiaire du schéma de généralisation. Si par exemple la pulsation de croches est immédiatement suivie d'une pulsation de noires, alors l'algorithme est en mesure de prolonger la pulsation de noires qui était initialement cachée au sein de la pulsation de croches.

### 5.1.4. Travaux similaires sur les problématiques d'ornementation.

L'approche géométrique de MEREDITH et al (2002) est appliquée directement à la polyphonie, et peut être utilisée dans le cadre de la recherche de motifs ornementés. Afin d'éviter l'extraction de motifs composés de notes excessivement distantes temporellement, une contrainte a été ajoutée basée sur une notion de compacité. Cette contrainte peut être reliée à la notion de relation syntagmatique entre notes temporellement proches.

L'approche par programmation dynamique proposée par ROLLAND (1999) est fondée sur un alignement de sous-séquences musicales utilisant des distances d'édition. Ces distances se fondent sur des opérations d'édition initialement définies pour le domaine du texte et de la biogénétique (insertion, suppression, remplacement), mais contiennent également quelques transformations spécifiques à l'expression musicale. Les opérations proposées dans notre approche permet un enrichissement de ces transformations par la prise en compte de configurations locales de notes au sein d'un espace paramétrique musical multidimensionnel.

### 5.1.5. Résultats actuels.

Nous avons tenté d'analyser, à l'aide de cette modélisation informatique, des productions musicales de styles variés, et notamment une improvisation modale arabe : *Istikhbar Mhayyer Sika*, jouée à la flûte *Nay* par Mohamed Saada (Figure 19), qui a été transcrite par Mondher Ayari. Les structures inférées par l'algorithme sont représentées par des lignes en dessous de chaque portée. Les motifs sont désignés par des symboles (1, 2, 3, 4, 5, + and -) à gauche de ces lignes. Les notes associées à chaque occurrence de motif sont représentées par des carrés alignés verticalement aux notes. Ces carrés représentent les états successifs au sein de la chaîne de motif, suivant une convention similaire à celle de la figure 3.

Le motif '-' représente une simple ligne mélodique conjointe descendante, c'est-à-dire une succession de hauteur de notes décroissantes. De manière analogue, le motif '+' représente une ligne mélodique conjointe ascendante. Les motifs 1 à 5 sont quant à eux répétés plusieurs fois au sein de l'improvisation. Les carrés noirs représentent le commencement d'une nouvelle occurrence, et les carrés blancs indiquent les états successifs au sein de chaque occurrence. Les carrés gris correspondent aux états optionnels qui n'existent pas au sein de la totalité des occurrences du motif. Enfin, les embranchements multiples indiquent une multiplicité de continuations possibles pour un même motif.

L'improvisation est construite autour du mode *Tba' Mhayyer Sika ré*, caractérisé par l'utilisation du mode (ré, mi, fa, sol, la, si bémol, do) et d'une figure mélodique spécifique, qui correspond exactement au motif 2. Le début de l'improvisation est construite autour d'une répétition successive du motif 1. Ce motif représente une courbe mélodique périodique, dont chaque nouvelle période commence et finit sur la même note fa, qui est donc une note-pivot de l'improvisation. La seconde ligne de l'improvisation est caractérisée par la répétition successive du motif 3, qui est une petite phrase mélodique transposée. Le motif 4 correspond à un autre profil mélodique associé au motif 2. Enfin, les deux dernières portées contiennent des répétitions du motif 5.

**Figure 19.** Analyse, à l'aide de *OMkanthus*, du début de l'improvisation de *Istikhbar Mhayyer Sika* jouée par Mohamed Saada.

## 5.2. Modélisation des mécanismes de segmentation

La phase suivante de la collaboration consiste en la modélisation de processus de segmentation et en la confrontation des segmentations produites par la modélisation avec celles qui sont mesurées par la procédure expérimentale auprès des auditeurs. La modélisation des stratégies de segmentation suit le cadre méthodologique défini précédemment : les segmentations fondées sur le principe de répétition motivique sont directement issues de l'algorithme d'extraction qui vient d'être présenté dans les section précédentes. Les autres stratégies de segmentation, par contre, nécessitent l'ajout de nouvelles heuristiques d'analyse musicale.

### 5.2.1. Engendrement d'une segmentation à partir de l'analyse motivique

L'algorithme développé dans notre approche met en évidence, pour chaque œuvre musicale considérée, un ensemble de motifs saillants déterminés en raison de leur répétition. A priori, aucune contrainte n'est imposée concernant les superpositions possibles entre différentes classes de motif : un

motif peut être inclus dans un autre motif (si toutefois il est « fermé » (cf. section 4)) et divers motifs peuvent s'enchevêtrer.

Cette logique fondée sur la recherche de répétitions n'explique qu'en partie le phénomène de perception de structures motiviques. En parallèle, une autre heuristique se fonde sur une représentation globale de l'ensemble des structures et sur la notion de *segmentation* : la découverte de structures motiviques répétées induit la mise en évidence de repères temporels au cours de l'évolution de l'œuvre musicale, lesquelles pouvant exercer une influence sur la recherche ultérieure de structures répétées. Toute régularité de points de segmentation engendre en particulier une attente de prolongement de cette régularité, les points de segmentations ainsi attendues peuvent modifier le processus d'extraction motivique.

Les points de segmentations sont engendrés directement à partir des structures extraites lors de l'analyse motivique, à l'aide de la méthode proposée par CAMBOUROPOULOS (2006), suivant laquelle le point de départ d'une structure motivique ou son point d'achèvement sont considérés comme des points de segmentation. Le résultat de la segmentation de l'improvisation de *Istikhbar Mhayyer Sika* à partir de l'analyse motivique présentée précédemment est présenté à la figure 20. Chaque segmentation est représentée par un trait vertical pointillé, au-dessous duquel sont indiqués le ou les motifs à l'origine de cette segmentation.

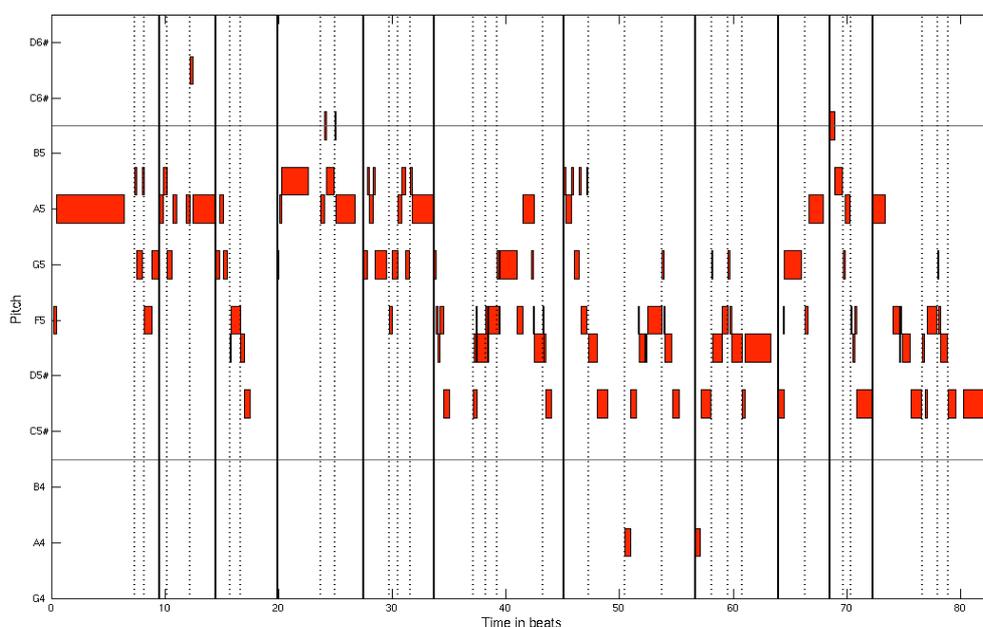
The figure shows a musical score for 'Nāy' in a single system with five staves. The score is written in a treble clef with a key signature of one flat (B-flat). The music consists of a single melodic line. Vertical dotted lines indicate segmentation points. Below each staff, numbers indicate the motifs that define the segmentation points. The motifs are: 1, 1, 2, 1, 2 (Staff 1); 3, 3, 3, 1,4, 2, 3, 1, 2,4 (Staff 2); 4, 2, 2,4, 2 (Staff 3); 5, 5 (Staff 4); 5, 1, 2, 5, 1, 2, 2, 4, 4 (Staff 5). Some motifs are grouped with brackets and a '3' above them, indicating triplets. The word 'Nāy' is written at the beginning of the first staff.

Figure 20. Traduction de l'analyse motivique de *Istikhbar Mhayyer Sika* en terme de segmentations.

### 5.2.2. Modéliser la segmentation par discontinuité auditive.

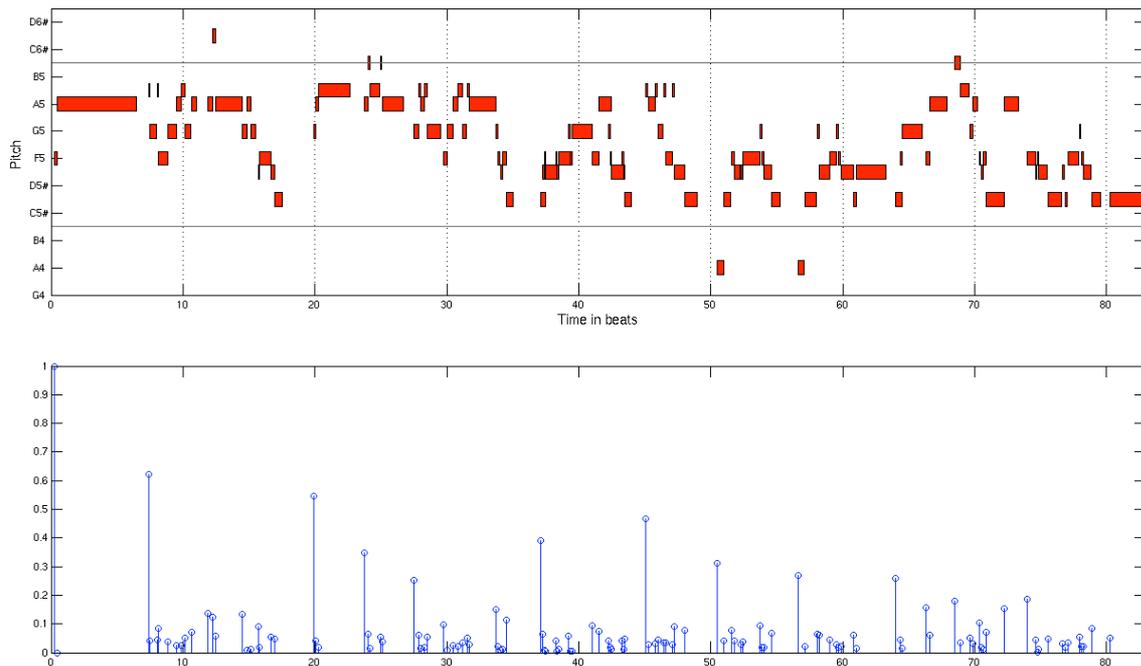
La modélisation de l'induction de structures musicales est fondée, jusqu'à présent, sur un mécanisme particulier : celui de recherche de répétitions. Sont ajoutés au modèle d'autres mécanismes, qui apportent une information complémentaire sur la structure musicale : l'un d'entre eux, en particulier, procède par agglomération de notes successives temporellement ou paramétriquement proches et, de manière duale, par disjonction de notes successives éloignées, suivant les principes *Gestalt* de proximité et similarité. Diverses modélisations du mécanisme d'agglomération/disjonction (Cf., par exemple, [TENNEY et POLANSKY 1980], [LERDAHL et JACKENDOFF 1983], [TEMPERLEY 1988], [NARMOUR 1990], [CAMBOUROPOULOS 1997, 2006]) ont été proposées. Certaines de ces approches ont été implémentées au sein de l'environnement informatique *MIDItoolbox* (EEROLA et TOIVIAINEN 2004), un module du logiciel *Matlab* développé au sein de l'Université de Jyväskylä.

La Figure 21 présente la segmentation de l'improvisation *Istikhbar Mhayyer Sika* à l'aide de l'algorithme de TENNEY et POLANSKY (1980). Le morceau est présenté sous la forme dite *piano-roll* : chaque note est indiquée par un rectangle, placé dans un espace bi-dimensionnel correspondant au déroulement temporel, en abscisse, et aux hauteurs de note, en ordonnée. Les propositions de segmentation sont indiquées par des traits verticaux, continus pour les segmentations principales et pointillés pour les segmentations secondaires.



**Figure 21.** Segmentation de l'improvisation *Istikhbar Mhayyer Sika* suivant la méthode de TENNEY et POLANSKY (1980).

La Figure 22 présente une autre hypothèse de segmentation proposée par la méthode de CAMBOUROPOULOS (2006). Les segmentations sont représentées sur le diagramme en bas de la figure. L'abscisse de ce diagramme correspond au même déroulement temporel que celui présenté dans le *piano-roll* en haut de la figure, l'ordonnée de ce diagramme représente le degré d'intensité des différentes hypothèses de segmentation : la valeur 1 étant associée à la segmentation d'intensité maximale, et les autres valeurs étant représentées de manière relative par rapport à cette valeur maximale.



**Figure 22.** Segmentation de l'improvisation *Istikhbar Mhayer Sika* suivant la méthode de CAMBOUROPOULOS (1997, 2006).

### 5.2.3. Intégration de schémas culturels.

La phase suivante du travail de modélisation est consacrée à l'articulation entre mécanismes de bas niveau et stratégies fondées sur des schémas culturels. La modélisation constituée lors de la première phase, bien qu'orientée vers les mécanismes de bas niveau, est implicitement influencée par le contexte culturel, lequel s'exprime à travers les corpus musicaux et les intuitions mises en oeuvre au cours de la démarche de modélisation. Une confrontation du modèle avec des corpus musicaux extra-européens a permis un remodelage des mécanismes de bas-niveau mais nécessite également une prise en compte des variantes interculturelles en termes de dépendance à des schémas culturels. L'intégration de schémas culturels au sein de la modélisation est envisagée sous deux formes : dans un premier temps, des schémas canoniques offerts par l'expertise ethno-musicologique, puis dans un second temps, des schémas résultants des analyses statistiques des tests d'écoute. Cette seconde phase de notre collaboration se poursuit actuellement en prolongement de l'ACI.

#### **5.2.4. Confrontation des résultats expérimentaux et informatiques.**

L'ensemble des stratégies de segmentations proposées aux précédents paragraphes peuvent alors être combinées afin de mettre en évidence les interactions possibles et d'en proposer une synthèse. La segmentation globale effectuée de manière algorithmique peut alors être comparée avec les segmentations mesurées lors des tests d'écoute. L'objectif étant une corrélation des prévisions calculées par la modélisation informatique avec les résultats mesurés auprès des auditeurs : la modélisation offrirait alors une explication hypothétique des stratégies sous-jacentes de segmentation sous la forme d'une description systématique et détaillée. La mise en adéquation de la modélisation informatique et des résultats d'écoute nécessite un affinage progressif de la modélisation, afin de déterminer les paramètres relatifs aux divers degrés de liberté qui restent à préciser. Par exemple, la sommation des diverses stratégies de segmentation requiert la détermination des coefficients affectés à chacun de ces facteurs. D'autre part, chaque stratégie peut également être susceptible d'éventuelles améliorations.

### **6. Autres axes de recherche concernant la modélisation.**

#### **6.1. Généraliser vers la polyphonie.**

La musique a été considérée, dans notre approche ainsi que dans la plupart des approches actuelles, sous la forme d'une séquence de notes successives, ou *monodies*. Mais de manière générale, la musique doit être envisagée en terme d'un flux de notes, une *polyphonie*, pouvant contenir en particulier des superpositions de notes, ou *accords*. Dans un tel flux peut être perçue, d'une part, une multiplicité de monodies simultanées. En prolongeant le formalisme issu de la théorie des graphes, on pourra dire qu'à partir du réseau initial que constitue le flux polyphonique émerge un graphe de chaînes monodiques, lequel pouvant présenter des configurations complexes telles que des embranchements et des entrelacements de chaînes. Suivant ces diverses chaînes monodiques peuvent alors être opérée la recherche de répétitions de motifs assurée par notre modélisation. Une généralisation de l'approche au cadre polyphonique nécessitera donc une formalisation et une modélisation de ce processus d'émergence du réseau monodique à partir du flux de notes polyphonique.

La difficulté dans cette problématique d'émergence de chaînes monodiques au sein d'un flux polyphonique est que, pour des exemples musicaux complexes, il peut parfois exister une grande combinatoire de solutions possibles, la préférence d'une solution particulière se fondant sur des critères subjectifs. Comme une étude de ce mécanisme isolé peut s'avérer problématique, on l'étudiera plutôt dans son interaction avec le mécanisme de recherche de répétitions motiviques. La tâche consistera alors à détecter des chaînes motiviques au sein du flux polyphonique. La recherche de répétition, qui, dans un cadre monodique, était opérée sur une simple chaîne monodique, doit ici être

envisagée suivant un parcours d'un réseau monodique construit de manière adaptative sur le flux polyphonique. La construction du réseau monodique s'effectuera en étroit couplage avec la recherche de répétition : lorsqu'une répétition motivique est en cours de découverte, les extensions successives du motif pourront être recherchées sur de plus nombreuses chaînes monodiques puisque des descriptions spécifiques sont ici attendues.

Au sein d'un flux polyphonique émergent non seulement des chaînes monodiques, mais également des groupements de notes qui s'agglomèrent en raison de leur proximité temporelle, que l'on pourra appeler des *agrégats*, et qui forme diverses constructions musicales telles que les accords, les degrés ou les modes. La problématique d'émergence d'agrégats, qui s'apparente à celle d'émergence de chaînes monodiques, devra être modélisée. On pourra alors tester l'hypothèse d'une influence de cette segmentation en termes d'agrégats sur la recherche de répétitions de motifs. La modélisation de cette hypothèse, laquelle est susceptible de jouer un rôle important dans l'analyse motivique, pourra être envisagée rapidement, avant même la construction du mécanisme d'émergence d'agrégats, par une simple intégration manuelle au sein de la description d'une séquence musicale de celle de ses agrégats.

En outre, de même que les motifs résultent de répétitions de chaînes monodiques, la répétition d'agrégats induit la formation de classes d'agrégats. Il sera souhaitable de rendre possible, dans cette modélisation, une détection de classes d'agrégats à partir de répétitions approximatives d'agrégats, présentant des notes supplémentaires ou manquantes. D'autre part, les agrégats se succédant forment eux-mêmes des chaînes d'agrégats. Là où la description de chaînes monodiques consistait simplement en une suite de distances entre notes successives, la description de chaînes d'agrégats consiste en un réseau complexe de distances entre notes d'agrégats successifs, ainsi que de distances entre classes d'agrégats, qu'il faudra formaliser et modéliser. Il s'agira par la suite de généraliser la problématique de recherche de répétition de chaînes monodiques aux chaînes d'agrégats, aboutissant ainsi à des motifs constitués d'agrégats. La résolution de l'ensemble de ces problèmes permettra d'assurer une modélisation complète des mécanismes en jeu dans l'analyse motivique.

La problématique de recherche de motifs polyphoniques a été très rarement considérée (Cf. [DOVEY 2001], [CONKLIN et ANAGNOSTOPOULOU 2001], [MEREDITH et al. 2002], [MEUDIC et SAINT-JAMES 2004]). L'approche de Meredith et al., en particulier, consiste en une détection des répétitions exactes de motifs exprimés sous la forme de coordonnées dans l'espace bi-dimensionnel du temps et du paramètre des hauteurs. En conséquence, cette approche ne peut pas découvrir des répétitions de motifs présentant des fluctuations le long de certaines des différentes dimensions musicales – transformation du rythme, ou des hauteurs, etc. –, lesquelles pourtant ne remettent pas en cause, chez les auditeurs, l'identification des motifs. L'approche envisagée ici permettra justement une telle identification adaptative le long des différentes dimensions, car la

démarche est fondée sur une modélisation la plus précise possible des stratégies d'écoute, et non sur une méthode fondée sur des critères purement géométriques et statistiques.

Pourra alors être entreprise une ouverture encore plus large du problème, en envisageant non seulement des motifs composés de notes et des motifs composés d'agrégats, mais également des motifs composés eux-même de motifs, configuration qui joue un rôle également important au sein de l'expression musicale. Certaines recherches cognitives ont proposé des formalisations de structures motiviques complexes [DEUTSCH et FEROE 1981], mais celles-ci se limitent au cas particulier de motifs composés d'une répétition multiple d'un seul motif, et aucune modélisation des mécanismes de découverte de telles structures n'y est associée. Nous entreprendrons cette modélisation et envisagerons de généraliser le problème, en n'imposant aucune contrainte particulière entre les motifs élémentaires successifs qui composent la structure motivique complexe. Se tissent alors des relations assez complexes, notamment entre les motifs élémentaires successifs, qu'il faudra donc formaliser puis modéliser [LARTILLOT et SAINT-JAMES 2005].

## **6.2. Les applications de l'outil informatique.**

Concernant l'informatique musicale, l'intérêt d'une telle démarche est multiple. Elle permet tout d'abord une analyse minutieuse du langage musical, rendant compte avec précision du foisonnement motivique, décelant les innombrables cellules musicales caractéristiques et leurs développements, et dégageant, *in fine*, une structure d'une grande complexité, qu'il serait ensuite possible d'appréhender selon diverses échelles d'observation (au niveau global, local, etc.). Les capacités combinatoires de l'informatique permettraient une unification des points de vue macroscopique (analyse de la forme) et microscopique (analyse des motifs) en une seule démarche tissant un vaste réseau de mises en correspondance. Les divers éléments décelés constituent les motifs, le réseau lui-même constitue la forme. De cette forme généralisée, il sera possible d'en effectuer une navigation, un parcours le long des différentes hiérarchies, permettant à l'utilisateur musicien de s'appropriier cette structure complexe et au chercheur d'explorer la nature du traitement des informations musicales selon la culture dont les connaissances sont incluses dans le modèle.

Dans un second temps, les mécanismes de découverte automatique de lois régissant le langage musical offriront une vision computationnelle des théories musicales traditionnelles, mais également une formalisation de musiques dont les théories sont partielles voire inexistantes. En effet, un système d'induction automatique suivant un modèle cognitif permettrait l'inférence de la grammaire cachée du langage sous-jacent à l'œuvre musicale. Ceci permettra d'une part de valider les modèles mis en jeu, et d'autre part d'offrir un outil d'analyse musicale automatique de grand intérêt pour la musicologie, pour les sciences cognitives et pour les nouvelles technologies. Enfin, le fruit de ces analyses, rendant compte de l'œuvre musicale avec une infime précision, pourra être utilisé dans un cadre compositionnel. Ce logiciel pourrait à long terme permettre une édition dynamique et interactive des

divers objets musicaux et ainsi la conception d'un nouvel outil d'aide à l'élaboration de nouvelles œuvres.

Une telle application des approches cognitives des mécanismes inductifs au domaine musical offrirait un cadre d'expérimentation idéal pour ce type de système complexe. En effet, contrairement aux précédents domaines d'applications tels que le langage, la représentation sémantique de l'environnement ou les théories scientifiques, la musique présente un cadre structurel exempt — du moins dans une première approche — de toute sémantique. La constitution du réseau conceptuel peut ainsi être entreprise uniquement à l'aide d'heuristiques générales régissant les mécanismes cognitifs élémentaires, sans la nécessité d'introduire un réseau conceptuel a priori.

### **6.3. Analyse de masses de données musicales**

Un important domaine d'application de l'automatisation de l'analyse motivique concerne l'analyse des masses de données musicales. Sur l'internet se développent en effet des immenses catalogues de musique en ligne, mises à la disposition du grand public, et qui occupent une part de marché de plus en plus dominante dans le commerce musical. La conception d'outils de gestion, de visualisation et de navigation au sein de ces bases de données, est une problématique de grande importance. Une communauté scientifique internationale s'est constituée autour de cette thématique, intitulée *Music Information Retrieval* [DOWNIE 2003], créant un lien étroit entre les recherches en informatique musicale et les applications industrielles.

Les travaux présentés dans cet article pourront être appliqués en particulier à deux problématiques majeures dans ce domaine : d'une part l'élaboration de mesures de distance de similarité entre œuvres musicales, et d'autre part le développement de méthodologies de recherche au sein de ces catalogues. La similarité musicale, tout d'abord, est envisagée dans les travaux récents à travers de multiples composantes de l'expression musicale, relative notamment à la texture sonore [AUCOUTURIER et PACHET 2002] D'autres recherches étudient au contraire l'impact des structures musicales plus profondes sur les jugements de similarité musicale [CRAWFORD et al 1998 ; LEMSTRÖM et al 2005]. La modélisation présentée ici, offrant une description précise de la dimension motivique de ces structures musicales, permettra en conséquence un approfondissement de ce champ de recherche, et pourra donc donner lieu à des applications intéressantes dans ce domaine.

Les systèmes de fouilles de données musicales, quant à elles, ont pour objectif de retrouver, au sein d'un catalogue, les œuvres correspondant à une requête spécifiée par l'utilisateur. Certaines approches dans ce domaine procèdent par un recoupement exhaustif de la requête avec chaque instant successif de chaque œuvre du catalogue musical [UKKONEN et al 2003]. À l'opposé, l'espace de recherche peut être réduit significativement par l'intermédiaire d'une analyse préalable de chaque œuvre du catalogue : la requête peut alors être simplement comparé avec les représentations compactes produites par les analyses préalables [HOOS et al 2001]. Cette focalisation sur les

structures les plus saillantes des œuvres musicales assure en outre une amélioration qualitative de la pertinence perceptive des résultats. Dans ce contexte, notre système d'analyse automatique offrira une description compacte de la composante motivique de chaque œuvre du catalogue, permettant à terme une amélioration des performances de fouilles de données.

Dans notre modélisation, les séquences musicales sont décrites suivant des paramètres symboliques directement issus du texte musical, c'est-à-dire de la partition. Les bases de données musicales en ligne, quant à elles, sont généralement décrites de manière soit semi-symbolique, soit sonore. La description semi-symbolique, tout d'abord, suit généralement le protocole MIDI de contrôle informatique des instruments de musique, dans laquelle les paramètres musicaux (hauteur et rythme) ne s'identifient pas exactement aux données musicales symboliques, mais en représentent une quantification numérique. Afin de pouvoir analyser de telles bases de données, il est nécessaire de mettre en œuvre une opération de transcription automatique des descriptions semi-symboliques en descriptions symboliques. Les méthodes proposées dans la littérature de quantification des hauteurs [CAMBOUROPOULOS 2003 ; CHEW et CHEN 2005, MEREDITH et WIGGINS 2005, TEMPERLEY 1988] et des durées [CEMGIL et KAPPEN 2003] ne répondent pas entièrement au problème, car cette opération de transcription se fonde partiellement sur la structuration motivique des œuvres, facteur qui a été peu pris en compte jusqu'alors. La modélisation de l'émergence de cette structuration offerte par notre approche permettra de prendre en compte cette interaction, afin d'envisager une amélioration des résultats.

Par contre, la majeure partie des bases de données musicales consiste en des enregistrements sonores. Une analyse de ces bases nécessite une opération de transcription automatique du sonore vers le symbolique, problématique d'une grande complexité et qui fait l'objet de nombreuses recherches actuelles. Sont proposés notamment des outils de transcription du sonore vers le semi-symbolique, que l'on pourra transcrire de nouveau vers une représentation symbolique. Mais ici encore, il semble exister une relation d'interdépendance entre l'opération de transcription à partir de la représentation sonore et la recherche de structuration musicale.

## **Conclusions**

Notre étude permet de mettre en évidence que les motifs musicaux résultent d'un ensemble de mécanismes interdépendants. Une automatisation de l'analyse motivique nécessite alors une modélisation précise de ces mécanismes, au sein d'un réseau conceptuel au sein duquel se tissent divers types de relations (telles que la relation de spécificité). Le mécanisme opératoire de la modélisation se décline sous la forme d'opérateurs de base appliqués à chaque phase successive de chaque structure. Nous avons tenté, petit à petit, de construire une telle modélisation. Ces travaux ont débuté à partir de certaines hypothèses de base, telles que l'identification de variations motiviques à l'aide de similarités numériques, qui se sont avérées, une fois modélisées, problématiques voire

fausses. La mise en place des mécanismes de base de découverte motivique a permis ensuite de mettre en évidence le phénomène de redondance combinatoire, dont il a été nécessaire de déceler les causes. Nous avons tenté alors de modéliser de la manière la plus simple possible les mécanismes permettant de gérer efficacement de telles redondances au sein d'un système cognitif. La stabilité du système total dépend de la bonne définition de chaque opérateur élémentaire : le moindre défaut pouvant engendrer un comportement chaotique et une explosion combinatoire dont il est parfois difficile d'en retrouver la cause. En raison de l'extrême difficulté de contrôler l'ensemble des mécanismes permettant d'assurer la pertinence des motifs découverts, on suppose qu'une modélisation capable d'offrir des résultats qui concordent avec les structures effectivement perçues présente un certain degré de véracité cognitive : un tel modèle pourrait être mis en œuvre de près ou de loin au sein du système cognitif humain. L'expérimentation informatique permet ainsi de réduire le champ des possibilités de la modélisation cognitive. Le modèle qui en résulte devra alors être soumis à une validation et un affinage par le biais de la psychologie expérimentale.

Les résultats actuels permettent de montrer l'intérêt d'une telle approche. D'une part, le système informatique qui résulte de cette étude offre au musicologue une description exhaustive des motifs d'une partition, qu'un auditeur ne peut découvrir seul que de manière partielle et non-explicite. L'intérêt direct de cet outil informatique consiste donc à offrir à l'auditeur le moyen d'approfondir son appréciation de l'œuvre musicale, en lui mettant en évidence l'ensemble de ces structures. Cette technologie donnera en outre à la musicologie les moyens d'ouvrir largement le champ des possibilités de l'analyse musicale, et répond ainsi à de nombreuses attentes (Cf. par exemple [RETI 1951], [MEYER 1956], [RUWET 1966], [NATTIEZ 1975], [LARTILLOT et SAINT-JAMES 2005]). D'autre part, l'expérimentation algorithmique permet de mettre en évidence la non-faisabilité de certaines hypothèses, et de proposer une modélisation qui offre un comportement satisfaisant et est susceptible d'être mis en œuvre, de près ou de loin, au sein du système cognitif humain. Enfin, une telle automatisation peut trouver des applications très intéressantes dans l'étude des bases de données musicales. Cette modélisation complexe introduit de nouvelles représentations et de nouveaux processus, qui enrichissent aussi bien le savoir musicologique que la compréhension cognitive. Une analyse automatique des bases de données musicales, en particulier à l'aide d'une modélisation par système complexe, pourra permettre d'une part aux utilisateurs et aux administrateurs de naviguer de manière intelligente au sein du réseau — en parcourant les œuvres musicales de même genre, de même style, etc. — et d'autre part d'offrir une vision nouvelle des œuvres musicales. En outre, de telles modélisations, si elles s'effectuent en collaboration avec des musicologues et des psychologues de la perception musicale, pourraient constituer une formalisation des théories musicales et psychologiques et fournir ainsi un outil de réflexion pour les sciences humaines consacrées aux phénomènes artistiques.

## Annexe A. Références

- [AGRAWAL et SKIRANT 1995] R. Agrawal et R. Skirant. Mining Sequential Patterns. *International Conference on Data Engineering*, Taipei, Taiwan, 1995.
- [ASSAYAG et al 1999] G. Assayag, C. Rueda, M. Laurson, C. Agon, O. Delerue : « Computer Assisted Composition at Ircam: From Patchwork to Openmusic », *Computer Music Journal*, 23(3), pp. 59-72, 1999.
- [AUCOUTURIER et PACHET 2002] J.-J. Aucouturier et F. Pachet. Music similarity measures : What's the use ? *International Conference on Music Information Retrieval*, 2002.
- [CAMBOUROPOULOS 1998] E. Cambouropoulos : *Towards a General Computational Theory of Musical Structures*, Thèse de l'Université de Edinbourg, 1998.
- [CAMBOUROPOULOS 2003] E. Cambouropoulos. Pitch spelling: A computational model. *Music Perception*, 20, p. 411-429, 2003.
- [CAMBOUROPOULOS 2006] E. Cambouropoulos : « Musical Parallelism and Melodic Segmentation: A Computational Approach », *Music Perception*, 23(3), pp. 249-268.
- [CEMGIL et KAPPEN 2003] A.T. Cemgil et B. Kappen. Monte Carlo methods for tempo tracking and rhythm quantization. *Journal of Artificial Intelligence Research*, 18, pp. 47-81, 2003.
- [CHEW et CHEN 2005] E. Chew et Y.C. Chen. Real-Time Pitch Spelling Using the Spiral Array. *Computer Music Journal*, 29(2), 2005.
- [CONKLIN 2002] D. Conklin. Representation and discovery of vertical patterns in music. Dans C. Anagnostopoulou, M. Ferrand et A. Smaill (éd.). *Music and Artificial Intelligence*, Springer-Verlag, pp. 32-42, 2002.
- [CONKLIN et ANAGNOSTOPOULOU 2001] Conklin, D., and C. Anagnostopoulou. 2001. "Representation and Discovery of Multiple Viewpoint Patterns." *Proceedings of the 2001 International Computer Music Conference*. San Francisco: International Computer Music Association, pp. 479-485.
- [CONKLIN et ANAGNOSTOPOULOU 2006] D. Conklin, C. Anagnostopoulou : « Discovery of segmental patterns in music », *INFORMS Journal on Computing*, 18(3), 2006.
- [COOK 1987] N. Cook : *A Guide to Musical Analysis*, J.M. Dent & Sons, 1987.
- [COPE 1996] D. Cope : *Experiments in Musical Intelligence*, A-R Editions, 1996.
- [CRAWFORD et al 1998] T. Crawford, C.S. Iliopoulos et R. Raman. String-matching techniques for musical similarity and melodic recognition. Dans W. Hewlett et E. Selfridge-Field (éd.). *Melodic similarity : Concepts, procedures and applications (Computing in Musicology 11)*, MIT Press, 1998.
- [DANNENBERG et HU 2002] R.B. Dannenberg et N. Hu. Pattern Discovery Techniques for Music Audio. *Journal of New Music Research*, 32(2), p. 153-164, 2002.
- [DANNENBERG et MONT-REYNAUD 1987] R. Dannenberg et B. Mont-Reynaud : « Following an Improvisation in Real-Time », *International Computer Music Conference*, 1987.
- [DELIÈGE 1987] I. Deliège : « Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules », *Music Perception*, 4(4), pp. 325-360, 1987.

[DESAIN et HONING 2001] P. Desain et H. Honing : « Quantization of musical time: a connectionist approach ». In Todd, P.M., and Loy, D.G. (éds.), *Music and Connectionism*, p. 150–167. MIT Press, 2001.

[DOVEY 2001] M.J. Dovey : « A technique for “regular expression” style searching in polyphonic music », *Proceedings of the 2<sup>nd</sup> International Conference on Music Information Retrieval*, 2001.

[DOWLING et FUJITANI 1971] W.J. Dowling et D.S. Fujitani : « Contour, interval, and pitch recognition in memory for melodies », *Journal of the Acoustical Society of America*, 49, 524–531, 1971.

[DOWLING et HARWOOD 1986] W.J. Dowling et D.L. Harwood : *Music Cognition*, Academic Press, 1986.

[DOWNIE 2003] J.S. Downie. Music information retrieval. *Annual Review of Information Science and Technology*, 37, p. 295-340, 2003.

[ECK et CASAGRANDE 2005] D. Eck et N. Casagrande : « Finding Meter in Music Using an Autocorrelation Phase Matrix and Shannon Entropy ». *Proceedings of the International Conference on Music Information Retrieval*, London, 2005.

[EDWORTHY 1985] J. Edworthy : « Melodic contour and musical structure ». In P. Howell, I. Cross and R. West, *Musical structure and cognition*. (chap. 7), Academic Press, 1985.

[EEROLA et TOIVIAINEN 2004] T. Eerola, P. Toivainen : « MIR in Matlab: The MIDI Toolbox », *Proceedings of the 5<sup>th</sup> International Conference on Music Information Retrieval*, pp. 22-27, 2004.

[FORTE 1977] A. Forte. *Structure of Atonal Music*. Yale University Press, 1977.

[GANTER et WILLE 1999] B. Ganter et R. Wille. *Formal Concept Analysis: Mathematical Foundations*, Springer-Verlag, 1999.

[HAN et al 1998] J. Han, W. Gong et Y. Yin. Mining Segment-Wise Periodic Patterns in Time-Related Databases. *International Conference on Knowledge Discovery and Data Mining*, 1998.

[HAN et al 1999] J. Han, G. Dong et Y. Yin. Efficient Mining of Partial Periodic Patterns in Time Series Database. *International Conference on Data Engineering*, 1999.

[HJELMSLEV 1966] L. Hjelmslev : *Prolégomènes à une théorie du langage*, 1966, traduction française : Éditions de Minuit, 1971.

[HOOS et al 2001] H. Hoos, K. Renz et M. Görg. GUIDO/MIR : an Experimental Musical Information Retrieval System based on GUIDO Music Notation. *International Conference on Music Information Retrieval*, 2001.

[HOWELL, CROSS et WEST 1985] P. Howell, I. Cross et R. West : *Musical Structure and Cognition*, Academic Press, 1985.

[HUSSLERL 1905] Edmund Husserl, *Leçons pour une phénoménologie de la conscience intime du temps*, 1905, traduction française, Presses Universitaire de France, 1964.

[LARGE et KOLEN 1994] E.W. Large et J.F. Kolen : « Resonance and the Perception of Musical Meter », *Connection Science* 6(2/3), 177-208, 1994.

[LARTILLOT 2004a] O. Lartillot : « A Musical Pattern Discovery System Founded on a Modeling of Listening Strategies », *Computer Music Journal* 28(3), pp. 53-67, 2004.

[**LARTILLOT 2004b**] O. Lartillot : « A Multi-Parametric and Redundancy-Filtering Approach to Pattern Identification », *Proceedings of the 5<sup>th</sup> International Conference on Music Information Retrieval*, pp. 132-137, 2004.

[**LARTILLOT 2004c**] O. Lartillot. *Fondements d'un système d'analyse musicale suivant une modélisation cognitive de l'écoute*. Thèse de doctorat, Université de Paris 6, 2004.

[**LARTILLOT 2005**] O. Lartillot. Multi-Dimensional Motivic Pattern Extraction Founded on Adaptive Redundancy Filtering. *Journal of New Music Research*, 34(4), pp. 375-393.

[**LARTILLOT et SAINT-JAMES 2005**] O. Lartillot, E. Saint-James : « Automating Motivic Analysis Through the Application of Perceptual Rules », *Music Query: Methods, Strategies, and User Studies (Computing in Musicology 13)*, MIT Press, à paraître.

[**LEMSTRÖM et al 2005**] K. Lemström, G. Navarro et Y. Pinzon. Bit-Parallel Algorithms for Transposition-Invariant Multi-Track String-Matching. *Journal of Discrete Algorithms* 3(2-4), p. 267-292, 2005.

[**LERDAHL et JACKENDOFF 1983**] F. Lerdahl et R. Jackendoff : *A Generative Theory of Tonal Music*, The MIT Press, 1983.

[**LÉVI-STRAUSS 1958**] C. Lévi-Strauss : *Anthropologie structurale*, Plon, 1958.

[**MA et HELLERSTEIN 2001**] S. Ma et J. Hellerstein. Mining partially periodic event patterns with unknown periods. *International Conference on Data Engineering*, 2001.

[**McADAMS et al 2004**] S. McAdams, G. Assayag, M. Ayari, O. Lartillot et C. Jambois : « Segmentation et reconnaissance de patterns dans la musique », *Actes des Journées Action Concertée Incitative (A.C.I.) « Systèmes Complexes en Sciences Sociales »*, Paris, 2004.

[**MEEK et BIRMINGHAM 2001**] C. Meek et W.P. Birmingham. Thematic Extractor. *International Conference on Music Information Retrieval*, 2001.

[**MEEÛS 1993**] N. MeeÛs : *Heinrich Schenker : Une introduction*, P : Pardaga, 1993.

[**MEREDITH et al 2002**] D. Meredith, K. Lemström, G.A. Wiggins : « Algorithms for Discovering Repeated Patterns in Multidimensional Representations of Polyphonic Music », *Journal of New Music Research*, 31(4), pp. 321-345, 2002.

[**MEREDITH et WIGGINS 2005**] D. Meredith et G. Wiggins. Comparing Pitch Spelling Algorithms. *International Conference on Music Information Retrieval*, London, 2005.

[**MEUDIC et SAINT-JAMES 2004**] B. Meudic, E. Saint-James : « Automatic Extraction of Approximate Repetitions in Polyphonic MIDI Files Based on Perceptive Criteria », in U.K. Will (éd.), *Computer Music Modelling and Retrieval*, Springer-Verlag, 2004.

[**MEYER 1956**] L.B. Meyer : *Emotion and Meaning in Music*, The University of Chicago Press, 1956.

[**NARMOUR 1990**] E. Narmour : *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*, The University of Chicago Press, 1990.

[**NATTIEZ 1975**] J.-J. Nattiez : *Fondements d'une sémiologie de la musique*, Union Générale d'Éditions, 1975.

[**RETI 1951**] R. Reti: *The Thematic Process in Music*, Macmillan Publishing, 1951.

- [**ROLLAND 1999**] P.-Y. Rolland : « Discovering Patterns in Musical Sequences », *Journal of New Music Research*, 28(4), pp. 335-350, 1999.
- [**ROUGET 1961**] G. Rouget: « Un chromatisme africain », *L'Homme*, I. 3, p. 41, 1961.
- [**ROWE 1993**] R. Rowe: *Interactive Music Systems (Machine Listening and Composing)*, The MIT Press, 1993.
- [**RUWET 1966**] N. Ruwet: « Méthodes d'analyse en musicologie », *Revue belge de musicologie*, 20, pp. 65-90, 1966.
- [**TANAKA et al. 2005**] Y. Tanaka, K. Iwamoto et K. Uehara. Discovery of Time-Series Motif from Multi-Dimensional Data Based on MDL Principle. *Machine Learning*, 58, 2005.
- [**TEMPERLEY 1988**] D. Temperley: *The Cognition of Basic Musical Structures*, MIT Press, 1988.
- [**TENNEY et POLANSKY 1980**] J. Tenney et L. Polansky : « Temporal Gestalt Perception in Music », *Journal of Music Theory*, 24, pp. 205-241, 1980.
- [**TOIVIAINEN 1998**] P. Toiviainen : « An Interactive MIDI Accompanist », *Computer Music Journal*, 22, (4), 63-75, 1998.
- [**TOIVIAINEN et EEROLA 2006**] P. Toiviainen et T. Eerola : « Autocorrelation in meter induction: The role of accent structure ». *Journal of Acoustical Society of America*, 119, (2), 1164-1170, 2006.
- [**UKKONEN et al 2003**] E. Ukkonen, K. Lemström et V. Mäkinen. Geometric Algorithms for Transposition Invariant Content-Based Music Retrieval. *International Conference on Music Information Retrieval*, p. 193-199, 2003.
- [**WHITE 1960**] B. White. « Recognition of distorted melodies », *American Journal of Psychology*, 73, p. 100-107, 1960.
- [**YANG et al. 2002**] J. Yang, W. Wang et P.S. Yu. InfoMiner+: Mining Partial Periodic Patterns with Gap Penalties. *IEEE International Conference on Data Mining*, 2002.
- [**ZAKI 2005**] M. Zaki. Efficient algorithms for mining closed itemsets and their lattice structure. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17 (4), p. 462-478, 2005.

## Annexe B. Exemples d'analyses motiviques.

### B.1. Exemple d'analyse : une Sonate de Mozart.

L'analyse du premier thème de la *Sonate en la majeur* K 331 de Mozart (Figure 23) met en évidence le motif élémentaire de base (a) répété et transposé, ainsi que la phrase de 4 mesures (b) répétée deux fois. Toutefois, en raison d'une légère transformation rythmique à la fin de la première occurrence, il n'est pas possible pour l'instant de détecter cette phrase dans sa totalité (b'). Est mis également en évidence la répétition successive d'un simple rythmique constitué d'une croche et d'une noire, formant ainsi une structure cyclique (c). Il est également intéressant de constater que le même algorithme capable de découvrir une répétition de phrase de 4 mesures détecte également des successions d'intervalles de seconde ascendantes  $diat = +1$  (e) ou descendantes  $diat = -1$  (d et f).

Figure 23. Analyse automatisée du premier thème de la *Sonate en la majeur* K 331 de Mozart.

### B.2. Analyse d'une chanson médiévale.

Notre modèle propose une analyse simple et assez pertinente d'un chant médiéval du XIV<sup>ème</sup> siècle, un *Geisslerlied* intitulé « *Maria muoter reinu mait* » (Figure 24). Il s'agit en fait du premier exemple d'application de la méthode d'analyse proposée par Nicolas Ruwet dans son article historique (RUWET, 1966). Il est à noter qu'en raison de la présence d'un suffixe du motif *B* précédant immédiatement les deux occurrences successives du motif entier, notre modèle unifie ces trois occurrences sous la forme d'un seul motif cyclique et ne peut donc pas, pour cette raison, délimiter une occurrence de *B* par portée, contrairement à ce qui est indiqué en grisé sur la figure.

The image displays a musical score analysis of the *Geisslerlied* titled « Maria muoter reinu mait ». The score is presented in four systems, each containing two staves. The first two systems are labeled 'A' and 'B' on the left and right respectively, indicating the two voices. The last two systems are labeled 'B' on the right, indicating the lower voice. The notation includes treble clefs, a key signature of one flat (B-flat), and a common time signature (C). Melodic motifs are identified and labeled with letters: 'e' for a cyclic motif consisting of an ascending line followed by a descending line; 'c' for a cyclic motif consisting of a single ascending line; and 'd' for a cyclic motif consisting of a single descending line. Brackets and lines connect these motifs across the staves to show their relationships and repetitions.

Figure 24. Analyse du *Geisslerlied* intitulé « Maria muoter reinu mait » légèrement simplifié par rapport à la version proposée dans (RUWET, 1966).

### B.3. Analyse d'une Invention de Bach.

L'algorithme a également été appliqué à l'analyse d'une *Invention* de Bach à deux voix. Figure 25 montre l'analyse des 14 premières mesures, et la table 1 présente la liste des motifs découverts. Le motif A est le thème de l'*Invention*, répété de manière alternative sur les deux voix. Ce thème est composé de deux parties : une ligne conjointe ascendante (motif cyclique *c*) et une ligne conjointe descendante (motif cyclique *d*), séparées par un intervalle de septième. Pour des raisons de clarté de présentation, ces structures détaillées ne sont mentionnées qu'au sein de la première occurrence du motif. Le thème original A, dans cet extrait, n'est pas transposé dans le domaine des classes de hauteurs, et peut donc être exprimé suivant cette dimension, comme indiqué dans la table sous la colonne intitulée « Description mélodique ». La description rythmique est identique pour chaque occurrence. Le motif est donc décrit le long de l'ensemble des dimensions rythmiques (pulsation et inter-onset), comme indiqué sous la colonne « Description rythmique ».

Une variation du thème, notée A', est répétée de manière successive dans la voix supérieure, puis dans la voix inférieure. Cette variation est due à une transformation du premier intervalle du motif, un intervalle de tierce mineure descendant au lieu d'un intervalle de seconde parfaite. Contrairement au motif original A, la variante A' est transposée, et est donc définie sous forme d'intervalle de hauteurs au lieu de la description originale en termes de hauteurs absolues.

Le thème A est prolongé d'un contre-sujet B, construit sur une succession de croches (motif cyclique *b*) et contenant la répétition d'une ligne ascendante (motif cyclique *e*). Ce contre-sujet,

lorsqu'il est joué dans la voix inférieure (mesures 5-6), contient de manière plus spécifique la répétition d'un motif formé d'intervalles de hauteurs (motif cyclique *i*).

Un autre motif *A''* apparaît par la suite, construit sur une succession de doubles-croches (motif cyclique *a*) formant deux lignes conjointes ascendantes (motif cyclique *h*). Les deux motifs cycliques *a* et *h* suggèrent deux stratégies distinctes de pulsation rythmique. Le motif *A''* est considéré comme une variation du début du thème original *A* : il contient la même structure rythmique, une même succession d'intervalles de hauteurs entre les trois premières notes, et une succession similaire d'intervalles de hauteurs entre les quatre dernières notes. La seconde voix inclut également une ligne motivique de basse (motif cyclique *C*) de croches.

L'analyse inclut également des structures motiviques offrant un intérêt musical et une saillance perceptive moindre. Le motif *D* construit autour d'une succession d'un intervalle de seconde descendante suivi d'une série d'intervalles de seconde ascendantes, n'est pas facilement détectable en raison de sa position faible au sein de la structure métrique, et la taille limitée de sa description en particulier. La seconde occurrence du motif *D* est également considérée comme le début d'une variante *A'''* du motif *A'*, qui ne présente pas, cette fois encore, de saillance forte, en raison de la position faible du motif au sein de la structure générale de la phrase mélodique. Le motif *E* présente des limitations similaires. Le filtrage des motifs *E* et *D* sera envisagé dans les travaux futurs par l'intégration d'une représentation métrique de plus haut niveau, laquelle indiquera la position de chaque motif au sein de chaque mesure et, dans certains cas, contraindra les occurrences de motifs à conserver une information métrique équivalente. Le motif cyclique *f*, quant à lui, résulte de la répétition successive d'un intervalle de quarte ascendante, mais contredit une fois de plus la structure métrique. Le motif cyclique *g* représente une simple oscillation locale entre contour ascendant et descendant, et peut être facilement filtré si nécessaire.

L'intégration du contour au sein de l'espace paramétrique entraîne, comme expliqué précédemment, une prolifération de structures motiviques présentant une saillance faible. L'heuristique présentée à la section 3.3. a été ici prise en compte. Celle-ci interdit l'identification suivant le paramètre de contour pour les premiers intervalles d'un motif, lesquels doivent être identifiés d'une manière plus spécifique ; seuls les intervalles ultérieurs du motif pourront être identifiés suivant le paramètre de contour. Cette heuristique permet un filtrage sélectif des motifs peu saillants, sans conséquence pour les motifs saillants.

Le modèle inclut également quelques descriptions redondantes d'autres motifs, incorrectement filtrées par l'algorithme. Cela est due en partie de défauts de modélisation à corriger par la suite ; cela suggère également, pour une autre part, la nécessaire intégration d'autres heuristiques de filtrage. Enfin, l'obtention de la représentation compacte présentée à la Figure 25 a nécessité une mise en ordre

manuelle des résultats engendrés par l’algorithme, mise en ordre qui devra être automatisée par la suite.

The figure shows a musical score for the first 21 measures of J.S. Bach's Invention in D minor, BWV 775. The score is divided into three systems, each with a treble and bass clef staff. Motif classes are labeled with letters and their constituent notes.

**System 1 (Measures 1-4):**

- Treble clef: Motif A (a-a-a, c-c-c) and Motif B (b-b-b, e-e-e, e'-f-f).
- Bass clef: Motif A (a-a-a, c-c-c).

**System 2 (Measures 5-8):**

- Treble clef: Motif A (e), Motif A' (g-g-g), and Motif A'' (E).
- Bass clef: Motif B (e'-e'-e'), Motif C (e'-e'-e'), and Motif C (b-b-b).

**System 3 (Measures 9-12):**

- Treble clef: Motif A'' (a-a-a, h-h-h) and Motif A''' (E, D).
- Bass clef: Motif A' (a-a-a, c-c-c) and Motif A' (a-a-a, c-c-c).

**Fig. 25** – Analyse motivique de l’Invention en ré mineur BWV 775 de J.S. Bach (21 premières mesures). Les occurrences de chaque classe de motifs sont représentées de manière distincte.

**Table 2.** Motives discovered in Bach's *Invention in D minor* BWV 775. For each motive, the melodic and rhythmic dimensions upon which its description is constructed is indicated. Motives *B* and *h* are heterogeneous: their melodic description is based on a succession of two different dimensions.

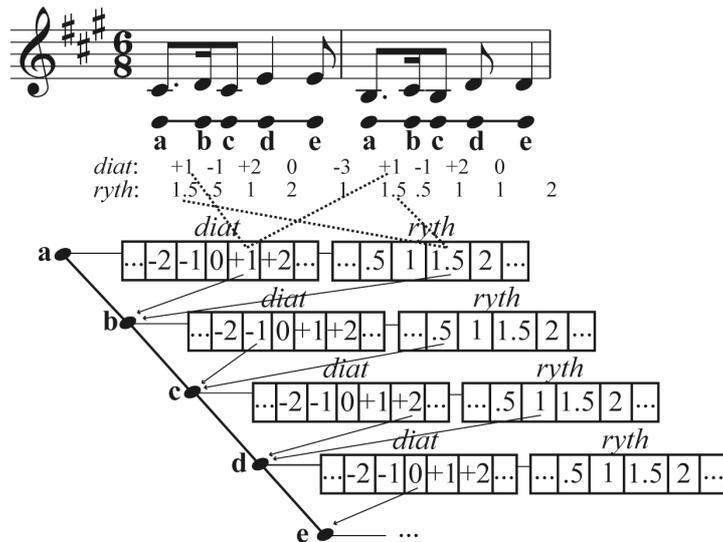
<b>Name</b>	<b>Function</b>	<b>Melodic description</b>	<b>Rhythmic description</b>
<b>A</b>	Theme	Pitch class	Complete
<b>A'</b>	Theme development	Pitch interval	Complete
<b>A''</b>	Cell developed from theme	Pitch interval	Complete
<b>B</b>	Countersubject	1st bar: pitch class 2nd bar: contour	Complete
<b>C</b>	Bass line	Pitch interval	Complete
<b>D</b>	<i>Artefact</i>	<i>Pitch interval</i>	<i>Complete except 1st interval</i>
<b>E</b>	<i>Artefact</i>	<i>Complete except 1st interval</i>	<i>Complete</i>
<b>a</b>	Series of 16th notes	No	Complete
<b>b</b>	Series of 8th notes	No	Complete
<b>c</b>	Ascending line	Pitch interval	Complete
<b>d</b>	Descending line	Pitch interval	Complete
<b>e</b>	Cell in countersubject	Contour	Complete
<b>e'</b>	Cell in 2nd countersubject	Pitch interval	Complete
<b>f</b>	<i>Increasing 4th interval</i>	<i>One pitch interval</i>	<i>Two rhythmic values</i>
<b>g</b>	<i>Increasing 2nd interval</i>	<i>One pitch interval</i>	<i>Two rhythmic values</i>
<b>h</b>	Sub-cell in cell A''	Absolute pitches and 1 last contour interval	Complete

## Annexe C. Détails concernant l'implémentation.

### C.1. L'algorithme de base de découverte motivique.

L'algorithme de découverte au cœur de la modélisation proposée se fonde sur le principe de *mémoire associative*, c'est-à-dire la capacité de mettre en relation des entités mémorisées présentant des propriétés communes. Cette mémoire associative est modélisée par l'intermédiaire de listes inversées associées à chaque paramètre musical (ici : mélodique et rythmique), et à chaque phase d'extension successive des motifs. Des premières tables permettent de mettre en relation les descriptions d'une œuvre musicale présentant un ou plusieurs paramètres en commun (FIG. 26). Ces descriptions sont alors reliées à un nouveau nœud de l'arbre, en descendance directe de la racine, représentant ce ou ces paramètres communs. Associées à ce nœud, de nouvelles tables mettent en relation, cette fois-ci, les descriptions qui succèdent aux descriptions précédentes, et qui présentent un ou plusieurs paramètres en commun. Ces nouvelles descriptions sont alors reliées à un nouveau nœud de l'arbre, en descendance directe du nœud précédent, et ainsi de suite. Ceci permet ainsi une découverte progressive des extensions successives de chaque motif. L'algorithme assure une découverte de motifs aussi bien homogènes qu'hétérogènes, car aucune contrainte n'est fixée *a priori* quant aux paramètres musicaux associés aux états successifs des motifs.

Un motif est découvert de manière incrémentale, par intégration progressive des intervalles successifs le constituant. D'abord est mise en évidence une répétition du premier intervalle du motif (intervalle *ab* sur la figure 5). Puis on découvre que certaines (ou l'ensemble) des occurrences de cet intervalle sont suivies d'un même intervalle, qui est alors ajouté au motif (devenant le motif *abc*), et ainsi de suite. Tous les intervalles sont mémorisés au sein de tables de hachage associées aux différents paramètres musicaux. En haut de la figure 26 sont ainsi représentées, à droite du nœud *a*, deux tables mémorisant les paramètres diatoniques (*diat*) et rythmiques (*ryth*) de l'ensemble des intervalles de la partition. Ces tables mettent en évidence par exemple une identité diatonique ( $diat = +1$ ) et rythmique ( $ryth = 1.5$ ) entre le premier intervalle de chaque mesure. Cette identité donne lieu à la création du motif *ab* associé à ces paramètres. Les intervalles succédant les occurrences du motif *ab* sont mémorisés au sein de nouvelles tables de hachage, à droite du nœud *b*. Ces tables mettent en évidence de nouveau une identité diatonique ( $diat = -1$ ) et rythmique ( $ryth = 0.5$ ) entre ces intervalles, laquelle donne lieu à l'extension du motif *ab* en un motif *abc*, et ainsi de suite. Les paramètres musicaux identifiés peuvent différer au fur et à mesure de l'extension progressive du motif. Par exemple, le dernier intervalle de chaque mesure ne s'identifie ici que suivant le paramètre diatonique ( $diat = 0$ ), car la valeur rythmique diffère : noire dans la première mesure ( $ryth = 2$ ), croche dans la seconde ( $ryth = 1$ ). Le dernier intervalle du motif *abcde* est donc associé uniquement à cette description diatonique.



**Figure 26.** Construction progressive du motif *abcde* (au-dessus de la partition) et de ses deux occurrences (sous la partition), par mémorisation des intervalles dans des tables de hachage associées à chaque phase successive du motif.

La mémoire associative *AssociativeMemory*(*Pattern*) associée à chaque motif *Pattern* est formalisée dans le pseudo-code listée à l'annexe D, sous la forme de listes inversées, qui associe à chaque valeur de chaque paramètre – ce qui est noté (*Parameter:Value*) – la liste des continuations possibles. Ces continuations sont représentées sous la forme (*Occurrence, Note*), où *Occurrence* du motif *Pattern* définit le contexte, avec la *Note* qui prolonge cette *Occurrence*, de telle manière que l'intervalle entre la fin de l'*Occurrence* et la nouvelle *Note* contient la valeur *Value* du paramètre *Parameter*.

### C.1.2. Découverte de motif

Des intervalles qui partagent une même valeur suivant une ou plusieurs dimensions musicales forment des occurrences d'un motif élémentaire, qui est représenté sous la forme d'une descendance directe (*b* à la figure) de la racine de l'arbre des motifs (*a*). La description du motif est la liste des identités paramétriques (indiquées par les flèches dans la figure) et la classe de motifs est l'ensemble des occurrences de ce motif. Chaque fois qu'un nouveau motif est créé, de nouvelles mémoires associatives (à droite du nœud *b*) mémorise tous les intervalles possibles qui suivent immédiatement les occurrences du nouveau motif (*b*). Lorsqu'une identité est détectée au sein de ces nouvelles tables, un nouveau motif est créé, en tant qu'extension du motif précédent (*c* extension du motif *b*), représenté comme descendant directe du nœud *b*, et ainsi de suite.

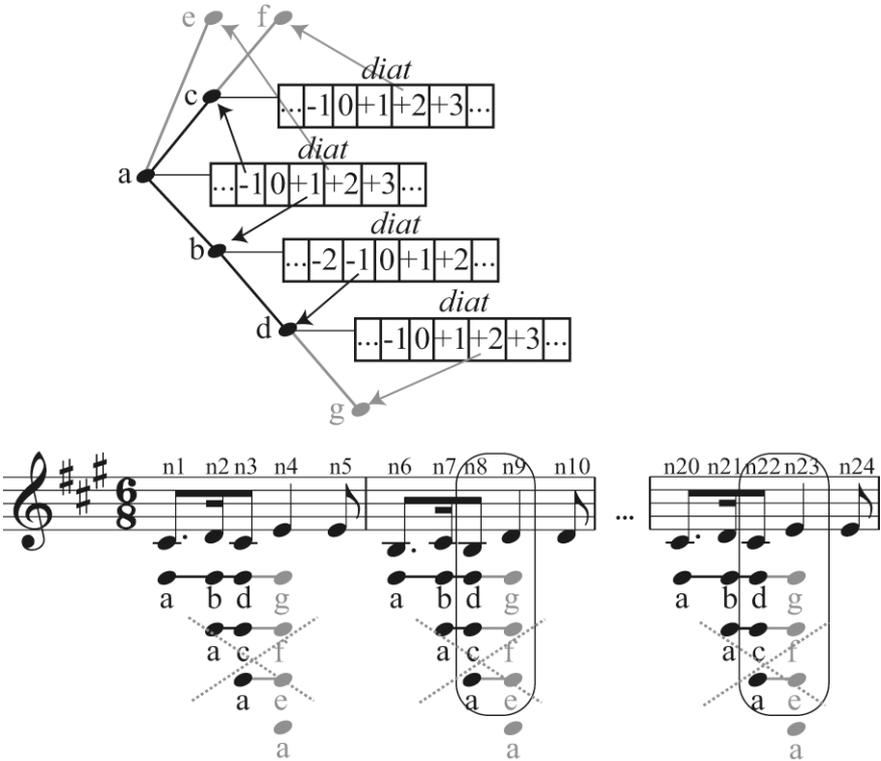
Cet algorithme permet une découverte progressive des extensions successives de chaque motif, homogènes ou hétérogènes (comme défini au paragraphe 2.2) : la sélection des dimensions

musicales définissant chaque extension successive du motif peut varier. Par exemple, à la figure, la dernière extension du motif *e* est simplement mélodique, car le rythme du dernier intervalle est différent pour chaque occurrence. D'autres contraintes ont été intégrées afin d'assurer un minimum de continuité entre les dimensions musicales des extensions successives des motifs.

Cette approche incrémentale permet une découverte multi-dimensionnelle adaptative de motifs. L'utilisation de listes inversées améliore les performances de calcul, car elles permettent un rappel direct d'anciens contextes similaires ne nécessitant aucune recherche au sein de la partition.

**Chronological Pattern Construction**

As justified in section 3.2.3, the pattern discovery process needs to be chronological: the main routine of the algorithm – called *ChronologicalPass* and described in the appendix – consists in a single pass through the musical sequence, from the first note  $n_i$  to the last note  $n_N$ . Each successive note launches a set of operations – listed in the function *AnalyseNewNote* described in the appendix – insuring the integration of the new note to the structures previously discovered and the creation of new structures. The detailed chain of operations contained in *AnalyseNewNote* is presented below.



**Figure 27.** Chronological analysis of the piece. When considering note  $n_9$ , patterns *e*, *f* and *g* are discovered. When considering note  $n_{23}$ , these patterns are simply recognized.

## Note Pattern Instantiation

As explained previously, each new note  $n_i$  is an occurrence of the note-pattern (represented by state  $a$  in figure 13). In this way, each note initiates its own pattern occurrence tree, which will contain all the possible occurrences starting from this note.

## Top-Down Traversal of the Specificity Graph

The interval between the two last notes  $n_{i-1}$  and  $n_i$  is constructed and is considered as a candidate extension of all the possible pattern occurrence chains concluded by the previous note  $n_{i-1}$ . We have seen that, following the closure principle, the selection of each pattern candidate depends on the support of their most specific patterns. Therefore the set of pattern occurrences concluded by the previous note  $n_{i-1}$  – which forms a graph called *Specificity Graph* – needs to be traversed in a decreasing order of specificity.

For instance, in figure 27, at step  $i = 9$ , patterns  $d$  and  $c$  can both be extended by notes  $n_9$  since the occurrences of both patterns are followed by the same description ( $diat=+2$ ). However, the resulting pattern  $f$  is non-closed since it has same support (2) than its super-pattern  $g$ , and should therefore not be created. That is why these patterns need to be considered in a decreasing order of specificity: first extending the most specific pattern  $d$  and then extending, if possible, the less specific pattern  $c$ .

It should be remarked however that, following our chronological approach, a new pattern class is inferred during the discovering of one of its occurrence. In this context, some of the more specific patterns do not apply to this occurrence at all, and should not be considered here. For instance, for a specific occurrence of a rhythmic pattern, the set of all possible melodic-rhythmic patterns of same rhythm should not be considered entirely. More precisely, the assessment of the redundancy of a candidate pattern simply depends on the pattern *occurrences* more specific than the occurrence currently considered and concluded by the same note.

For each successive pattern occurrence  $P_{i-1}$  (and in parallel each corresponding pattern  $P$  of which  $P_{i-1}$  is an occurrence) found during this traversal (at step  $i = 9$  in Figure 27, successively occurrences of  $d$ ,  $c$  and  $a$ ), two tests are operated, namely pattern *discovery* and pattern *recognition*, corresponding to the two subroutines *Discover* and *Recognise* described in the appendix. But first of all, the interval  $n_{i-1} n_i$  ( $n_8 n_9$  in figure 27) is memorised in the associative memories related to the candidate patterns ( $d$ ,  $c$  and  $a$ ), as described in the function *Memorise* in the appendix.

## Pattern Discovery

As explained in paragraph 4.1.2, any identity along one or several parameters implies the possible extension of pattern  $P$  with a new child  $D$  associated with this identified description. In figure 13, for instance, at note  $n_9$ , pattern  $c$  can be extended into  $f$  since the previous occurrence at note  $n_3$  accepts the same extension along the description  $diat = +2$ . The closure condition should however apply: there should not exist a more specific pattern  $R$  of same support. In figure 27, the extension  $f$  is in fact aborted since the more specific pattern  $g$  has same support (2). The set of more specific patterns  $R$  is constructed by extending the corresponding set of patterns  $Q$  more specific than the original pattern  $P$  by the same element  $n_i$ . Hence pattern  $g$  is retrieved as an extension of pattern  $d$ , which is more specific than  $c$ .

A pattern that is considered as redundant at one moment of the musical sequence may become non-redundant once it appears alone, without the more specific description, at a later stage of the analysis. For instance, in Figure 7, when only the two first occurrences are analysed, both patterns having same support, only the more specific pattern  $e$  should be explicitly represented. But the less specific pattern  $i$  will be represented once the last occurrence is discovered, as it is not an occurrence of the more specific pattern  $e$ .

It should be remarked that, on the other hand, even when a pattern  $x$  is considered as a non-redundant suffix of another pattern  $y$ , its extension  $x'$  may, on the contrary, become redundant. This happens when the pattern class of  $x'$  is strictly included into the pattern class of  $x$  and becomes equal to the pattern class of a more specific pattern  $y'$ . For this reason, the redundancy of a pattern should be checked at every phase of its extension.

Each update of the pattern tree by the extension of a pattern  $P$  by a new child  $D$  (such as the extension of  $d$  into  $g$  in figure 27) is associated with an update of the pattern occurrence trees: the current pattern occurrence chain  $P_{i-1}$  (occurrence of  $d$  concluded by note  $n_8$ ) is extended with a new child  $D_i$  (child  $g$  associated with note  $n_9$ ), occurrence of  $D$ . Moreover, all the other previous pattern occurrence chain  $P_{j-1}$  are also extended with a new child  $D_j$ , every time the interval  $n_{j-1} n_j$  complies with the description of  $D_i$ . For instance, in figure 27, the occurrence of pattern  $d$  concluded by note  $n_3$  is extended with a child  $g$  associated with note  $n_4$ .

When the following note  $n_{i+1}$  ( $n_{10}$ ) will be considered, the new interval  $n_i n_{i+1}$  ( $n_9 n_{10}$ ) will be automatically memorised in the associative memory of the new patterns  $D$  ( $g$ ) induced by the previous note  $n_i$ . The interval  $n_j n_{j+1}$  ( $n_4 n_5$ ), on the contrary, could not be stored in the same way, since, at the moment this interval was considered, the pattern  $D$  was not already discovered. Therefore, each time a

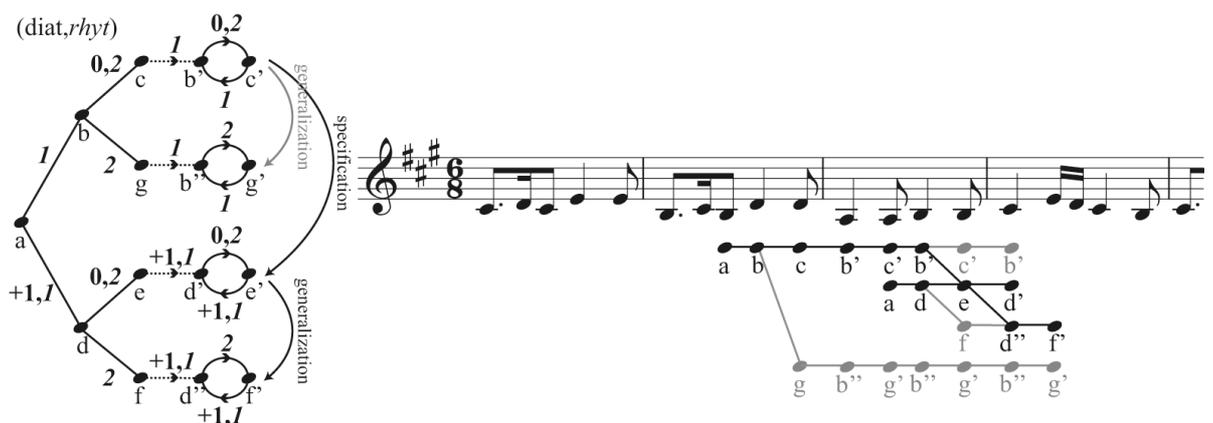


black, and should not therefore be represented explicitly. However, the generalization of pattern  $i$  (bar 6) into pattern  $j$  (bar 7) leads to the implicit generalization of pattern  $c$  into pattern  $k$ , than can therefore be immediately identified in bar 8.

The less specific patterns, although not explicitly represented in the analysis, should be updated if necessary. In particular, when a pattern is generalised, the more general patterns should be generalised too. For instance, as  $i$  has been generalised into  $j$ ,  $c$  should also be generalised into  $k$  in the same way. Hence, at next bar (8) the general pattern  $k$  is simply recognised.

### General and specific cycles.

The specificity relation defined in paragraph 3.1 had to be applied to cyclic patterns too: a cyclic pattern  $C$  would be considered as more specific than another cyclic pattern  $D$  when the sequence of description of pattern  $D$  is included in the sequence of descriptions of pattern  $C$ . Here too, this concept of specificity plays a pivotal role in music perception and enables a sound algorithmic processing of music. In Figure 29, the seven first notes of the cycle oscillate around the cyclic PC  $b'-c'$  simply composed of an oscillation between two rhythmic values of 8<sup>th</sup> note and quarter note ( $rhyt = 1$  and 2), the second interval also associated with a unison interval ( $diat = 0$ ). Then a more specific cycle  $d'-e'$  includes an ascending interval ( $diat = +1$ ), and is generalised after four notes into cycle  $d''-f''$  that does not feature the unison interval any more. Moreover, following the rule of generalisation of generalised patterns explained in previous paragraph, the more general cycle  $b'-c'$  needs to be generalised into a cycle  $b''-g''$  where the unison interval has been discarded. The integration of this phenomenon into the model helps insuring the relevance of the results and avoiding numerous unwanted combinatorial redundancies.



**Figure 29.** Actually, complex cyclic patterns are perceived: a first oscillation between a 8th note and a quarter note associated with a unison interval ( $b'-c'$ ) is specified through the integration of an ascending interval associated with the 8th note ( $d'-e'$ ). The cycle is then generalized due to the absence of the unison ( $d''-f''$ ) and cycle  $a'-b'$  is also generalized, for the same reason, into  $b''-g''$ .

## D. Pseudo-Code des mécanismes de base de la modélisation

### **ChronologicalPass:**

```
FOR each successive Note of the musical sequence
    CALL AnalyseNewNote(Note)
ENDFOR
```

### **AnalyseNewNote (Note) :**

```
CREATE an occurrence of NotePattern related to Note
FOR each Pattern Occurrence concluded by the previous note, from the most
specific to the less specific ones
    CALL Memorise(Pattern, Occurrence, Note)
    CALL Recognise(Pattern, Occurrence, Note)
    CALL Discover(Pattern, Occurrence, Note)
END FOR
```

### **Memorise (Pattern, Occurrence, Note) :**

```
FOR each musical Parameter
    LET Value be the Parameter value related to Note
    ADD the new context (Occurrence, Note) to the
AssociativeMemory(Pattern) at corresponding (Parameter:Value) address
END FOR
```

### **Recognise (Pattern, Occurrence, Note) :**

```
FOR each Extension of Pattern
    IF Note complies with the Extension description
        CREATE a NewOccurrence of Extension by extending Occurrence
with Note
        CALL DetectCyclicity(NewOccurrence)
    END IF
END FOR
```

### **Discover (Pattern, Occurrence, Note) :**

```
FOR each possible Description of Note, from the most specific to the less
specific ones
    IF Description is not included in an existing extension of Pattern
        LET Contexts be the intersection of the lists returned by
AssociativeMemory(Pattern) for each (Parameter:Value) contained in
Description
        REMOVE from Contexts any couple (RecalledOccurrence,
RecalledPattern) that does not confirm the condition
FigureGroundRule(RecalledOccurrence, RecalledPattern)
        IF there remains Contexts older than current context
            AND IF the number of Contexts is higher than the class of each
pattern more specific than (Pattern, Description)
                CREATE a NewPattern as an extension of Pattern with a new
state related to Description
                FOR each (RecalledOccurrence, RecalledNote) in Contexts
                    CREATE a NewOccurrence of NewPattern by extending
RecalledOccurrence with RecalledNote
                    CALL DetectCyclicity(NewOccurrence)
                END FOR
            END IF
        END IF
    END FOR
```

### **DetectCyclicity (NewOccurrence) :**

```
LET Pattern be the pattern associated to NewOccurrence
IF Pattern is not cyclic
AND IF NewOccurrence is immediately preceded by an occurrence of Pattern or
a pattern more specific than Pattern
    CREATE a CyclicPattern related to Pattern
```

END IF

**FigureGroundRule(Occurrence,Note) :**

LET PreviousNote be the note that concludes Occurrence  
THERE IS NO OtherOccurrence concluded by PreviousNote  
    SUCH THAT the pattern of OtherOccurrence be cyclic  
    AND Occurrence be more specific than the OtherOccurrence  
    AND OtherOccurrence be longer than Occurrence  
    AND the description of the extension of the OtherOccurrence be more  
specific than the description of Note