

Du son au sens : vers une approche sub-symbolique de l'analyse musicale assistée par ordinateur

Musurgia, vol. XVIII, n° 1-2, p. 100-116.

Résumé :

L'utilisation de l'ordinateur a donné lieu à deux types d'approche, l'une consistant à manipuler et modéliser des informations sonores et musicales, l'autre exploitant les possibilités multimodales. L'analyse computationnelle s'est naturellement orientée vers le paradigme symbolique qui, appliqué à la musique, consiste à analyser les éléments de la partition codés symboliquement (notes, accords, durées, rythmes, nuances, etc). Nous présentons une nouvelle approche de l'AMAO fondée sur le paradigme sub-symbolique. L'extraction d'informations sub-symboliques opérée par des descripteurs psychoacoustiques offre la possibilité d'approcher la perception de plusieurs dimensions musicales telles que le temps (tempo, attaques, durées), la hauteur (rugosité, saillance, multiplicité, clarté tonale, etc.) et le timbre (brillance, inharmonicité, enveloppe, clarté spectrale, etc.). Ces outils d'analyses sont particulièrement adaptés aux musiques du XX^e siècle et apportent des solutions originales pour l'analyse des processus à grande échelle. Afin d'illustrer l'utilisation des descripteurs psychoacoustiques quelques-uns d'entre eux sont employés pour l'analyse de la *Sequenza VII* de Luciano Berio. Ils ont permis de répondre à des questions précises sur la répartition des attaques en fonction du canevas temporel de la pièce, sur la saillance perceptive des classes de hauteurs et sur les phénomènes de tension/détente produits par la combinaison des dissonances sensorielle et harmonique.

Du son au sens : vers une approche sub-symbolique de l'analyse musicale assistée par ordinateur

Philippe LALITTE*

Introduction

L'ordinateur possède de nombreux atouts pour l'analyse. D'une part, ses capacités de stockage, d'indexation et de calcul en font un outil précieux pour traiter des grandes quantités de données. D'autre part, ses interfaces fournissent des fonctionnalités de visualisation, de lecture et de synchronisation entre son et image. De fait, computation et multimédia sont à la base du développement de l'analyse musicale assistée par ordinateur (AMAO) depuis une quarantaine d'années. L'utilisation de l'ordinateur a, en effet, donné lieu à deux types d'approche, l'une consistant à manipuler et modéliser des informations sonores et musicales, l'autre exploitant les possibilités multimodales. Après avoir présenté succinctement ces deux tendances de l'AMAO, nous développerons quelques arguments en faveur d'une approche fondée sur le paradigme sub-symbolique. Nous terminerons par quelques exemples d'utilisation de descripteurs psychoacoustiques appliqués à l'analyse de la *Sequenza VII* de Luciano Berio.

État de l'art de l'analyse musicale assistée par ordinateur

Bien que l'on puisse faire remonter l'origine de l'AMAO aux premières recherches en musique algorithmique – celles de Lejaren A. Hiller et Leonard M. Isaacson aux Etats-Unis ou celles de Pierre Barbaud¹ en France –, il faut cependant attendre les années 1980 pour voir apparaître les premiers outils informatiques réellement

* Maître de conférences à l'Université de Bourgogne, Centre Georges Chevrier (UMR CNRS 5605) et Laboratoire d'étude de l'apprentissage et du développement (LEAD CNRS UMR 5022), philippe.lalitte@u-bourgogne.fr.

¹ Pierre BARBAUD, *La musique, discipline scientifique*, Paris, Dunod, 1968.

dédiés à l'analyse musicale. L'approche computationnelle de l'analyse musicale a déjà une riche histoire dont il est impossible de rendre compte de façon exhaustive dans le cadre de cet article. Les recherches ont porté autant sur la description et la manipulation de contenu musical, que sur la simulation de styles ou la segmentation et la reconnaissance automatiques de *patterns*. En France André Riotte et Marcel Mesnage font figure de pionniers avec le Morphoscope² qui, tout en permettant la visualisation de la partition et sa segmentation, offre de nombreuses fonctions d'analyses à partir de composantes définies par l'utilisateur. Il existe actuellement plusieurs logiciels qui permettent d'effectuer des comptages, des classements, des comparaisons de données musicales. Certains d'entre eux fonctionnent à partir d'un codage spécifique comme Humdrum de David Huron³, d'autres recourent à la norme MIDI comme Monika, le programme de description de monodies (avec une résolution en quart de tons) de Nicolas Meeùs⁴, ou la Midi Toolbox (fonctionnant sous Matlab) développée par Petri Toiviainen et Tuomas Eerola⁵. Les données « brutes » obtenues par ces outils doivent ensuite être analysées en fonction du contexte esthétique et viennent la plupart du temps en complément de méthodes classiques d'analyse.

Quelques théories d'analyse ont été transposées dans l'univers informatique. Ainsi, la théorie de la ligne fondamentale d'Heinrich Schenker a fait (et fait encore) l'objet de plusieurs adaptations informatiques (Michael Kessler⁶, Stephen Smoliar⁷). La Set Theory d'Allen Forte a connu plusieurs implémentations, notamment celle de Peter Castine incluant la fonction intervallique de David Lewin et le vecteur d'intervalles adjacents de Richard Chrisman⁸. La bibliothèque *Dn* d'OpenMusic, développée par Moreno Andreatta et Carlos Agon, intègre certains paradigmes algébriques de la Set Theory. Il faut mentionner que le Morphoscope et OpenMusic⁹ ont aussi été utilisés comme outil de modélisation de processus

² Le Morphoscope et Musinote ont ensuite fusionné dans un nouveau logiciel Musiscope. Un certain nombre d'analyses réalisées avec le Morphoscope peuvent être consultées dans André RIOTTE et Marcel MESNAGE, *Formalismes et modèles musicaux, un recueil de textes (1963-1998)*, vol. 2 Exemples de modélisations de partitions musicales, Paris, Ircam-Centre Pompidou / Editions Delatour, 2006.

³ David HURON, « Humdrum and Kern: Selective feature encoding », Eleanor Selfridge-Field (éd.), *Beyond MIDI: The Handbook of Musical Codes*, Cambridge, the MIT Press, 1997, p.375-401.

⁴ <http://www.plm.paris-sorbonne.fr/spip.php?article213>.

⁵ Petri TOIVIAINEN et Tuomas EEROLA, « MIR in Matlab: The Midi Toolbox », *Proceedings of 5th International Conference on Music Information Retrieval*, Barcelone, Université de Pompeu Fabra, 2004, p. 22-27.

⁶ Michael KASSLER, « Transferring a Tonality Theory to a Computer », Daniel Hearz et Bonnie C. Wade (éd.), *International Musicological Society: Report of the Twelfth Congress, Berkeley 1977*, Kassel, Bärenreiter, 1981, p. 339-347.

⁷ Stephen W. SMOLIAR, « A Computer Aid for Schenkerian Analysis », *Computer Music Journal* 4 (1980), p. 41-59.

⁸ Peter CASTINE, *Set Theory Objects. Abstractions for Computer-Aided Analysis and Composition of Serial and Atonal Music*, Francfort, Peter Lang, 1994.

⁹ Carlos AGON, Moreno ANDREATTA, G. ASSAYAG, Stéphane SCHAUB, « Formal Aspects of Iannis Xenakis' Symbolic Music: A Computer-Aided Exploration of Compositional Processes », *Journal of New Music Research* 33/2 (2004).

compositionnels. David Temperley¹⁰ a fondé son approche sur la *Théorie générative de la musique tonale* de Lerdahl et Jackendoff. Temperley modélise le rythme, les frontières de groupement, les fonctions harmoniques et la séparation des voix d'une polyphonie en y associant des facteurs de pondération. Certaines recherches ont préféré se focaliser sur la simulation de styles musicaux spécifiques. Ainsi, le logiciel EMI¹¹ de David Cope produit des simulations assez réalistes de différents styles musicaux (Bach, Mozart, Beethoven) en exploitant des fonctions d'apprentissage de similarités. Mario Baroni, Maurizio Jacoboni et Rossana Dalmonde ont utilisé le principe de la grammaire générative pour modéliser le style de Legrenzi et l'implémenter dans le logiciel LEGRE¹².

D'autres chercheurs ont proposé des systèmes fondés sur des modèles cognitifs (*Gestalttheorie*, induction, similarité) et des techniques d'apprentissage, moins dépendants de théories analytiques ou de styles particuliers. Ainsi, l'algorithme de détection des frontières locales (*Unscramble*) d'Emilios Cambouropoulos¹³ s'articule sur les règles de changement d'identité et de proximité¹⁴. Plusieurs bibliothèques d'OpenMusic sont dédiées à la segmentation et à la reconnaissance automatiques de *patterns* (*Morphologie* de Jacopo Baboni-Schilingi et Frédéric Voisin, *Kanthume* d'Olivier Lartillot, *OMkant* et *MusicMap* de Benoît Meudic). Alors que *Kanthume* inclut la taille de la mémoire à court terme, *MusicMap* utilise des matrices de similarité.

L'interface graphique de l'ordinateur offre plusieurs possibilités de notation musicale (partition, tablature, piano-roll du séquenceur MIDI, etc.) et de visualisation du son (forme d'onde et sonagramme). Les musicologues ont rapidement compris l'intérêt de bénéficier de divers types de représentations selon le contexte musical. L'usage du sonagramme¹⁵ remonte aux recherches sur la parole menées aux Etats-Unis à la fin des années 1930. Emile Leipp en acoustique et Tran Van Khê en ethnomusicologie¹⁶ ont repris cette technique en France dans les années 1960. L'analyse sonographique permet de relever des régularités dans le continuum sonore, de repérer les attaques, de dissocier des phénomènes sonores complexes et, selon François Delalande, d'effectuer « toutes sortes de réductions, de

¹⁰ David TEMPERLEY, *The Cognition of Basic Musical Structures*, Cambridge, The MIT Press, 2001.

¹¹ David COPE, *Experiments in Musical Intelligence*, Madison, A-R Editions, 1996.

¹² Mario BARONI, Rossana DALMONDE et Carlo JACOBONI, *Le Regole della Musica*, Turin, EDT, 1999.

¹³ Emilios CAMBOUROPOULOS, « A Formal Theory for the Discovery of Local Boundaries in a Melodic Surface », Gérard Assayag, Marc Chemillier et Jean-Claude Eloy (éd.), *Troisièmes journées d'informatique musicale JIM 96*, Les cahiers du GREYC, vol. 4, 1996.

¹⁴ Emilios CAMBOUROPOULOS et Pierre-Yves ROLLAND, « Algorithmes et techniques pour l'analyse musicale », *Informatique musicale : du signal au signe musical*, François Pachet et Jean-Pierre Briot (éd.), Paris, Lavoisier, 2004, p. 332.

¹⁵ Analyse à partir d'un sonagramme où l'axe vertical représente les fréquences et l'axe horizontal le temps ; l'intensité des composantes étant représentée sous la forme de nuances de gris.

¹⁶ Tran Van KHÊ, « L'utilisation du sonographe dans l'étude du rythme », *Revue de musicologie* 54 (1968), p. 222-232.

transcriptions, de mesures »¹⁷. Dans son ouvrage *New Images of Musical Sound*¹⁸, Robert Cogan est l'un des premiers à procéder à des analyses, à partir d'images sonographiques, d'œuvres de styles et d'époques variés (de la musique tibétaine au jazz). Depuis, l'analyse musicale fondée sur l'étude du sonagramme a conduit à des méthodologies spécifiques comme la spectromorphologie¹⁹ de Denis Smalley, destinée essentiellement à l'analyse de la musique électroacoustique. Le sonagramme est employé, aujourd'hui, dans divers contextes historiques et stylistiques comme en témoignent l'étude du rondeau *Rose, Lis* de Guillaume de Machaut par Daniel Leech-Wilkinson²⁰ ou celle d'une chanson d'Elvis Costello par David Brackett²¹. Signalons qu'il existe aussi le cochléogramme, une représentation sonographique fondée, non sur le signal acoustique, mais sur une image « perceptive » simulant la traduction du signal dans la cochlée²².

La possibilité de synchroniser l'image et le son a ouvert de nouvelles voies pour l'AMAO. L'Acousmographe, développé par l'INA-GRM (O. Koechlin, H. Vinet et D. Bultiauw)²³, a été conçu comme un outil d'analyse auditive et de transcription. Il opère le passage de la visualisation sonographique vers la transcription en objets graphiques. À la possibilité de synchronisation, le logiciel ajoute des fonctions de paramétrage et des outils destinés à décrire le contenu musical à l'aide de textes ou de graphiques (ceux-ci pouvant se substituer au sonagramme). Plusieurs exemples d'analyses réalisées avec l'Acousmographe sont consultables sur le site du GRM (*Portraits Polychromes*)²⁴. Le logiciel iAnalyse²⁵ de Pierre Couprie se situe dans la continuité de l'Acousmographe, mais à partir d'une partition synchronisée à un fichier audio. Outre l'annotation textuelle et graphique, le logiciel se caractérise par « l'utilisation d'une base de données dans laquelle est stocké l'ensemble des informations entrées par l'utilisateur et permettant ainsi de réaliser des représentations différentes à partir d'une même analyse, l'usage de véritables outils d'aide à l'analyse musicale (les fonctions musicales) et la possibilité d'exporter

¹⁷ François DELALANDE, « L'articulation interne/externe et la détermination des pertinences en analyse », Jean-Marc Chouvel et Fabien Levy (éd.), *Observation, analyse, modèle. Peut-on parler d'art avec les outils de la science ?*, Paris, L'Harmattan / Ircam / Centre Georges Pompidou, 2002, p. 181.

¹⁸ Robert COGAN, *New Images of Musical Sound*, Cambridge, Harvard University Press, 1984.

¹⁹ Denis SMALLEY, « Spectromorphology and Structuring Processes », Simon Emmerson (éd.), *The Language of Electroacoustic Music*, New York, Palgrave MacMillan, 1986, p.61-93.

²⁰ Daniel LEECH-WILKINSON, « *Rose, lis* revisité », Jacqueline Cerquiglini-Toulet et Nigel Wilkins (éd.), *Guillaume de Machaut, 1300-2000*, Paris, PUPS, 2002, p. 53-69. [« Musiques-Écritures, Série Études »]

²¹ David BRACKETT, *Interpreting Popular Music*, Berkeley, University of California Press, 2000.

²² David M. HOWARD et Andy M. TYRELL, « Psychoacoustically informed spectrography and timbre », *Organised Sound* 2 (1997), p. 65-76.

²³ <http://www.ina.fr/entreprise/activites/recherches-musicales/acousmographe.html> (consulté le 19 mars 2009).

²⁴ <http://www.ina.fr/entreprise/activites/recherches-musicales/portraits-polychromes.html> (consulté le 19 mars 2009).

²⁵ <http://logiciels.pierrecouprie.fr/spip.php?rubrique1> (consulté le 19 mars 2009).

l'analyse dans différents formats »²⁶. D'autres technologies ont été développées pour l'analyse multimédia suscitant actuellement une réflexion et un développement sans précédent. Les « clefs d'écoute » du laboratoire d'ethnomusicologie du Musée de l'Homme à Paris²⁷, les « Ecoutes signées »²⁸ et le logiciel Music Lab 2 Maquette²⁹ de l'Ircam, les « Guides d'écoute » de la médiathèque de la Cité de la musique³⁰, les recherches du MIM sur les Unités sémiotiques temporelles³¹, les articles publiés en ligne par les revues Musimédiane³² et Déméter³³, constituent autant de directions fructueuses pour l'analyse musicale en elle-même, mais aussi pour son enseignement et sa réception.

Le paradigme sub-symbolique

On considère généralement, en sciences cognitives, qu'il existe deux principaux paradigmes pour modéliser le traitement des informations par le cerveau : le paradigme symbolique (à la base du cognitivisme) et le paradigme sub-symbolique (lié au connexionnisme). Le premier envisage le système cognitif comme un système de traitement de l'information qui manipule des symboles par le biais de règles syntaxiques. Chaque symbole possède sa propre représentation mentale mémorisée. Pour le second, il n'existe pas de représentation en soi, mais seulement des représentations distribuées sous forme de connexions neuronales. Dans l'approche connexionniste, les représentations sont moins stables, mais plus en interaction avec l'environnement. Ces deux paradigmes, qui ont donné lieu à des débats passionnés, tendent aujourd'hui à se fondre dans une vision intégrative.

Appliqué à la musique, le paradigme symbolique consiste à analyser les éléments de la partition codés symboliquement (notes, accords, durées, rythmes, nuances, etc.) et les règles syntaxiques les organisant. La plupart des théories musicologiques relèvent de ce paradigme et considèrent la note comme l'unité musicale fondamentale. Ainsi Jean-Jacques Nattiez affirme que, quelles que soient les cultures, « l'unité minimale dont est faite une pièce musicale est la note, ou, plus exactement, l'unité scalaire »³⁴. Dans le paradigme sub-symbolique, les unités minimales ne sont pas elles-mêmes des représentations. Celles-ci émergent de la

²⁶ Pierre COUPRIE, « iAnalyse : un logiciel d'aide à l'analyse musicale », *Journées d'informatique musicale*, Albi, 27-29 mars 2008, <http://www.gmea.net/jim08/index.php/Articles>, Consulté le 19 mars 2009.

²⁷ <http://recherche.ircam.fr/equipes/repmus/marc/ethnomus/ecoute/index.html> (consulté le 19 mars 2009).

²⁸ http://apm.ircam.fr/performance_analysis/ (consulté le 19 mars 2009).

²⁹ http://sticf.univ-lemans.fr/num/vol2006/bresson-07/sticf_2006_bresson_07p.pdf (consulté le 19 mars 2009).

³⁰ <http://www.cite-musique.fr> (consulté le 19 mars 2009).

³¹ <http://www.labo-mim.org/site/index.php?2008/08/11/24-les-ust> (consulté le 19 mars 2009).

³² <http://www.musimediane.com/> (consulté le 19 mars 2009).

³³ <http://demeter.revue.univ-lille3.fr/> (consulté le 19 mars 2009).

³⁴ Jean-Jacques NATTIEZ, « La signification comme paramètre musicale », *Musiques Une encyclopédie pour le XXI^e siècle*, Jean-Jacques Nattiez (dir.), vol. 2, *Les savoirs musicaux*, Paris, Actes Sud / Cité de la Musique, 2004, p. 258.

détection d'invariants et de caractéristiques distinctes contenus dans l'environnement. Les mécanismes de bas niveau (transduction, analyse et traitement par le système auditif) jouent alors un rôle prépondérant. Il est donc possible, selon Marc Leman, d'imaginer que « la connexion entre une représentation mentale et le monde réel peut être fourni par un modèle des mécanismes perceptifs sensoriels et de bas niveau du système auditif »³⁵. Nous sommes donc en présence de deux façons de penser les phénomènes musicaux. D'une part, une démarche descendante (*top-down*), qui pense le musical d'abord en termes de symboles traduits ultérieurement en sons et d'autre part une démarche ascendante (*bottom-up*) qui, du niveau sensoriel, conduit à la signification.

L'analyse computationnelle s'est naturellement orientée vers le paradigme symbolique. En effet, les données symboliques réduisent la réalité sonore et sont par conséquent facilement manipulables par l'ordinateur. Qui plus est, la manipulation de symboles s'inscrit dans la continuité des méthodes d'analyse développées par la musicologie systématique. L'abstraction inhérente aux données symboliques a permis d'élaborer de nombreuses méthodes d'analyse computationnelle pour modéliser les structures de hauteurs et de durées. Cependant, lorsqu'il s'agit de modéliser des dimensions musicales moins facilement réductibles, le paradigme symbolique semble marquer le pas (au moins jusqu'à présent). Ainsi, le codage MIDI n'offre qu'une approximation de la dynamique et du timbre. Certaines informations comme les changements continus (variations de timbre lors d'un crescendo par exemple, fluctuations de hauteur) ou les variations fines de tempo, sont moins facilement transposables dans le domaine symbolique. Dans le champ de la musique contemporaine, les méthodes computationnelles travaillant avec des données symboliques permettent également de modéliser les aspects de hauteur et de rythme, mais se heurtent à la question du timbre. Dans le cas des musiques instrumentales, la notation des modes de jeu et les indications d'instrumentation n'ont souvent qu'un faible pouvoir prédictif et ne permettent pas d'analyse systématique du timbre. La notation traditionnelle devient même totalement obsolète, pour l'analyse, lorsqu'elle ne sert qu'à indiquer à l'interprète les déclenchements de sons d'un échantillonneur³⁶.

Le recours au sonagramme est venu pallier dans une certaine mesure les déficiences de la notation traditionnelle, voire l'absence de notation dans le cas de la musique contemporaine (qu'il s'agisse de musique instrumentale ou de musique électroacoustique). Cependant, la pertinence du sonagramme n'est pas inhérente à l'outil. Une utilisation adéquate nécessite des réglages effectués en amont que ce soit la résolution en fréquence de l'analyse ou les paramètres de visualisation (résolution en temps et en fréquence, luminosité, contraste, couleurs, etc.). Jean-Claude Risset a signalé un certain nombre d'autres problèmes liés à ce type de

³⁵ Marc LEMAN, *Music and Schema Theory, Cognitive Foundations of Systematic Musicology*, Berlin, Springer, 1995, p. 181-182 ; traduit par mes soins.

³⁶ C'est le cas de nombreuses œuvres mixtes qui utilisent un échantillonneur. La hauteur du son déclenché peut être totalement différente de la note écrite.

représentation comme « les différences entre un son harmonique et un son légèrement inharmonique, ou entre deux relations d'intervalle, l'une consonante, l'autre dissonante »³⁷. On peut mentionner également certains phénomènes difficilement repérables comme la fusion auditive, due par exemple à l'harmonicité, à la cohérence de la modulation d'amplitude, ou à la corrélation temporelle des attaques, etc. Ces constatations nous orientent vers une nouvelle méthode d'AMAO fondée sur le paradigme sub-symbolique mettant en œuvre des descripteurs psychoacoustiques.

Les descripteurs psychoacoustiques extraient d'un signal audio un certain nombre d'informations sub-symboliques. Ils opèrent une réduction du signal³⁸ qui permet de modéliser des attributs perceptifs tels que la hauteur, la rugosité, le brillance du timbre, le tempo, voire des aspects de plus haut niveau comme la tonalité ou la structure formelle d'une pièce musicale. À l'origine, les descripteurs psychoacoustiques ont été conçus dans le domaine de l'analyse de la parole. Ils sont particulièrement exploités, aujourd'hui, dans le domaine de la recherche d'information musicale (*Music Information Retrieval*)³⁹ pour l'indexation dans des bases de données. D'autres domaines de recherches comme les sciences de la perception musicale⁴⁰ (psychoacoustique, psychologie cognitive, neurosciences) ou l'analyse de l'interprétation exploitent les descripteurs psychoacoustiques. Plusieurs logiciels utilisables dans le cadre de l'analyse musicale existent actuellement comme l'IPEM toolbox (IPEM Gand)⁴¹, Psysound 3 (University of Sydney et University of New South Wales)⁴² ou Sonic visualiser (Queen Mary University, Londres)⁴³. Le potentiel de ces outils informatiques semble prometteur. Nous avons tenté une première approche de leur utilisation, à laquelle nous renvoyons le lecteur, pour expliciter la relation d'harmonie/timbre chez Varèse⁴⁴. Nous analyserons ici trois aspects de la *Sequenza VII* de Luciano Berio afin d'illustrer quelques-unes des possibilités des descripteurs psychoacoustiques.

³⁷ Jean-Claude RISSET, « Problèmes posés par l'analyse d'œuvres musicales dont la réalisation fait appel à l'informatique », *Analyse et création musicale : actes du troisième congrès européen d'analyse musicale*, Paris, L'Harmattan, 2001, p. 141-142.

³⁸ Alors que le fichier audio contient le signal dans son intégralité (à l'échantillonnage et à la quantification près), le fichier de description ne contient qu'un type d'information (par exemple, les variations dans le temps de *loudness*).

³⁹ Une des principales questions du domaine concerne l'estimation de la similarité entre des pièces musicales stockées sous forme de fichiers audio.

⁴⁰ Voir par exemple Philippe LALITTE, Emmanuel BIGAND, Joanna KANTOR-MARTYRUSKA et Charles DELBÉ, « On listening to atonal variants of two Piano Sonatas by Beethoven », *Music Perception* 26/3 (2009), p. 223-234.

⁴¹ <http://www.ipem.ugent.be/Toolbox/> (consulté le 19 mars 2009).

⁴² <http://psysound.wikidot.com/> (consulté le 19 mars 2009).

⁴³ <http://www.sonicvisualiser.org/> (consulté le 19 mars 2009).

⁴⁴ Philippe LALITTE, « Son organisé et spéculation sur les distances chez Varèse », *Cahiers de la Société Québécoise de Recherche Musicale* 9/1-2 (2007), p. 125-139. Philippe LALITTE, « Densité 21,5, un condensé d'harmonie-timbre », T. Horodyski et P. Lalitte (éd.), *Varèse : du son organisé aux arts audio*, actes des journées Varèse, 30 et 31 mars 2006, Université Paris 8, Paris, L'Harmattan, 2007, p. 245-273.

Exemples d'application à partir de la *Sequenza VII* de Luciano Berio

Il ne s'agit pas de fournir, dans le cadre de cet article, une analyse détaillée de l'œuvre. Notre travail se situe plutôt dans une optique exploratoire afin d'examiner dans quelle mesure les descripteurs psychoacoustiques peuvent être utiles à l'analyse musicale. La *Sequenza VII*⁴⁵ étant l'une des pièces les plus connues du répertoire contemporain pour hautbois, nous ne présenterons que les aspects essentiels à la compréhension de l'analyse qui va suivre⁴⁶. Selon Berio, la *Sequenza VII* « suggère deux types d'écoute : une écoute objective qui suit les événements dans leur ordre d'apparition et une écoute analytique à la fois timbrale, harmonique et polyphonique qui attrape les relations à distance, la polyphonie cachée, la perspective spatiale centrée autour de *si* ... »⁴⁷. La pièce joue effectivement sur la tension entre des événements produits à une courte échelle temporelle (notes répétées, gestes, figures) et de ceux se produisant à une échelle beaucoup plus large (déploiement du champ harmonique, densité de notes, variations de timbre, etc.). Dans les quelques exemples qui vont suivre, nous nous attacherons au deuxième type d'événements car les descripteurs psychoacoustiques semblent particulièrement pertinents pour mettre en évidence les processus à grande échelle.

Densité des attaques

La partition de la *Sequenza VII* est écrite en notation proportionnelle⁴⁸. Elle se présente sous la forme d'une seule grande feuille au format A2 comportant 13 lignes (systèmes), chacune étant divisée par 13 colonnes délimitant 169 mesures⁴⁹. Une durée chronométrique est assignée à chaque colonne de sorte que les durées des mesures décroissent dans chaque ligne (durées en secondes : 3 / 2,7 / 2 / 2 / 2 / 2 / 1,8 / 1,5 / 1,3 / 1,3 / 1 / 1 / 1). Ce canevas temporel devrait provoquer – en théorie – une densité de notes plus importante en début de ligne qu'en fin de ligne (plus il y a de temps à l'intérieur d'une mesure, plus le nombre potentiel de notes est élevé). On peut donc se demander si la contrainte temporelle imposée par ce canevas a influencé l'écriture de Berio. Une réponse possible à cette question consiste à mesurer la densité d'attaques (*onset*) avec un algorithme de détection automatique⁵⁰. La figure 1 représente la moyenne des attaques par ligne et par colonne. La moyenne des attaques par colonne dégage un profil décroissant, ce qui correspond globalement à la décroissance des durées de chaque colonne. De ce

⁴⁵ Luciano BERIO, *Sequenza VII per oboe solo* (1969), London, Universal Edition, 1971, UE 13754.

⁴⁶ Nous renvoyons le lecteur à l'analyse de Pierre Michel dans Jean-Yves BOSSEUR et Pierre MICHEL, *Perspectives analytiques (1950-1985)*, Paris, Minerve, 2007, p. 266-274.

⁴⁷ Luciano BERIO, « Chemins en musique », *La Revue musicale* 375-376-377 (1985), p. 438.

⁴⁸ Une version en notation mesurée a été réalisée en 2005 en collaboration avec l'hautboïste Jacqueline Leclair.

⁴⁹ Le choix du chiffre 13 aurait été dicté par le nombre de lettres du nom et du prénom du dédicataire, comme d'ailleurs celui de la note pivot *si* (« h » en notation allemande).

⁵⁰ Nous avons utilisé l'un des algorithmes du logiciel Sonic Visualiser. L'enregistrement provient de la version de Lazlo HADADI, *Berio: Sequenzas*, Deutsche Grammophon, 1998, CD 457 038-2.

point de vue, Berio a respecté la contrainte de la grille. Par contre, l'observation des moyennes d'attaques par ligne montre un profil beaucoup plus instable : croissance jusqu'à ligne 7, stabilisation à un niveau plus faible lignes 8 et 9, croissance lignes 10 et 11, décroissance lignes 12 et 13. Berio obéit dans ce cas à une logique harmonique. La densité d'attaques semble suivre le remplissage de l'espace des hauteurs (figure 3) : accroissement de la densité d'attaque jusqu'à ce que onze classes de hauteurs aient été énoncées, stabilisation à l'apparition de la douzième classe de hauteur (*sol*), et ensuite décroissance. Il est intéressant de noter que la chute de la densité d'attaques lors des deux dernières lignes de la partition correspond à l'apparition massive de points d'orgue et des sons multiphoniques. Les deux processus (temporel et harmonique) superposés forment un déploiement complexe de distribution des attaques représenté dans la figure 2. L'effet de périodicité auquel on aurait pu s'attendre — une vague de notes déferlant au début de chaque ligne — est atténué par la dramaturgie imposée par le déploiement harmonique. Quelques pics (entre 14 et 21 attaques par mesure) apparaissent, le plus souvent en début de ligne (L3/C1 ; L3/C7 ; L7/C2 ; L7/C7 ; L9/C1 ; L10/C5 ; L11/C1 ; L11/C4-5-6 ; L12/C1-2-3). Le regroupement de quelques pics dans le premier tiers des lignes 11 et 12 suggère une intensification de la vélocité lors de ces deux passages.

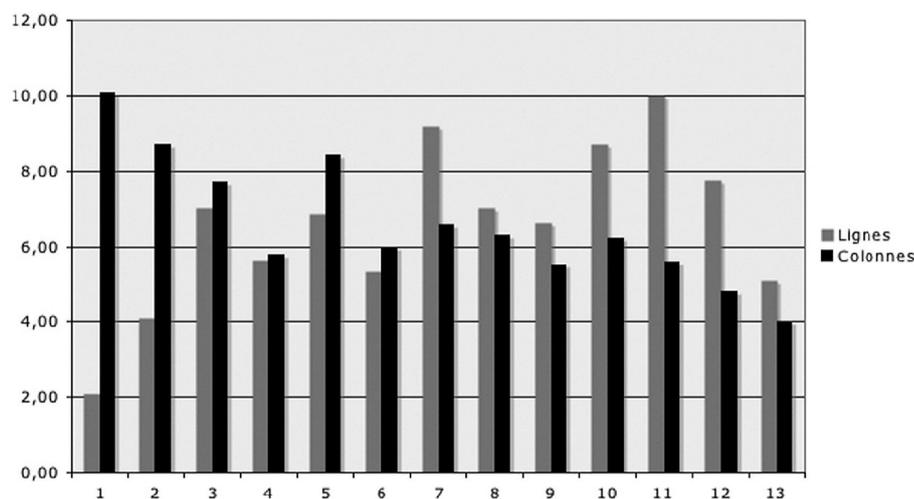


Figure 1 : Moyennes des densités d'attaques par ligne et par colonne

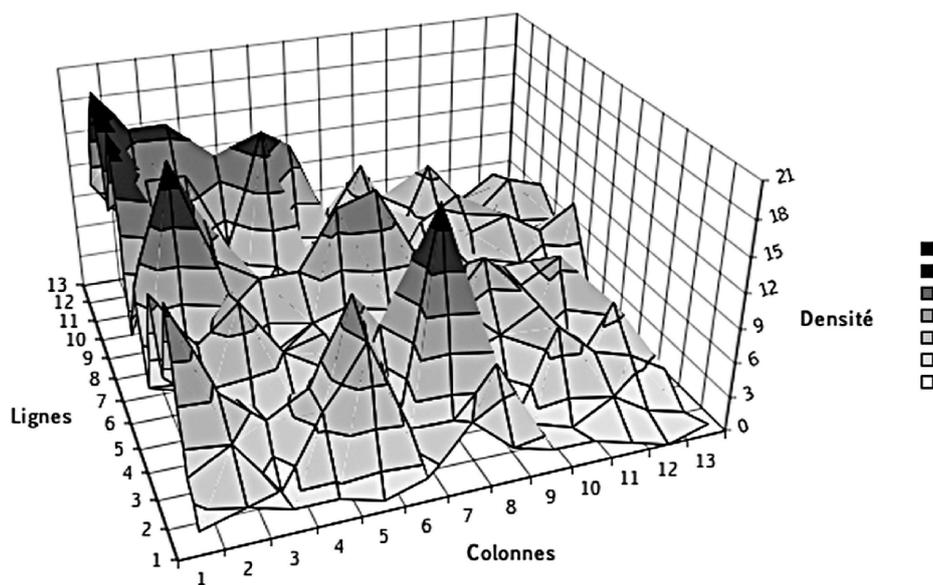


Figure 2 : Représentation en 3D des densités d'attaques (*onset*)

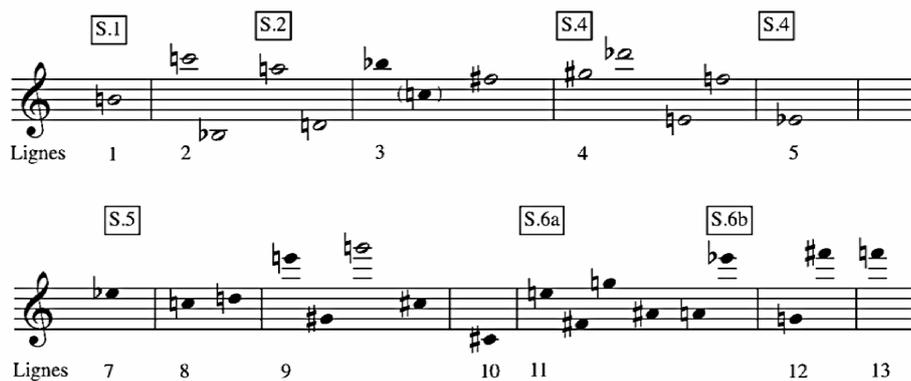


Figure 3 : Ordre d'apparition des classes de hauteurs (notes rondes) et des hauteurs (notes blanches et noires) ; les chiffres indiquent quand apparaissent pour la première fois les hauteurs ; les encadrés indiquent les sections

Saillance des classes de hauteurs

D'un point de vue harmonique, les hauteurs de la *Sequenza VII* se déploient par un remplissage progressif de l'espace contenu par l'ambitus de la pièce (*si 2/sol5*), celui-ci correspondant quasiment à l'étendue de l'instrument (*si 2/la5*). La figure 3 montre l'ordre d'apparition des classes de hauteurs (notes blanches) et des hauteurs (notes blanches et noires). Seules cinq hauteurs entre *si 2* et *sol5* sont absentes de la partition (*si2/do3/fa3/si4/ré5*). Le point le plus important concerne la progression qui, du *si3*, aboutit à la culmination sur *sol5* (ligne 9). Le compositeur a lui-même souligné l'importance de cette organisation des hauteurs et des registres : « Une fois que la dernière de ces douze notes *sol* est apparue, la pièce est presque terminée... »⁵¹.

L'ensemble des variations de timbre, cumulé aux variations de dynamiques, d'attaques et de durées, tend à bouleverser la perception des hauteurs. Pierre Michel a souligné la difficulté, avec cette œuvre, « de ne penser qu'en termes de hauteurs tant elles sont tributaires des sonorités obtenues par les transformations dues aux doigtés (l'intonation en est parfois légèrement modifiée aussi) »⁵². Il est donc légitime de se poser la question de la perceptibilité du déploiement harmonique prévu par Berio. Comment mesurer le poids perceptif de chaque classe de hauteurs ? Suivant l'approche symbolique, il suffit de calculer le nombre d'occurrences d'une classe de hauteurs dans la pièce entière ou dans chaque mesure. Cependant, si l'on veut tenir compte des variations de timbre pour analyser l'importance structurelle des classes de hauteurs, il vaut mieux se fier au paradigme sub-symbolique. Nous avons utilisé un descripteur de la saillance perceptive des classes de hauteurs (*chroma salience*)⁵³. L'algorithme mesure le poids de chacune des douze classes de hauteurs selon le modèle de Richard Parncutt⁵⁴. Le tableau suivant donne les valeurs de saillance pour chaque classe de hauteurs. Logiquement la saillance de *si* (1.06) est nettement plus forte que celle des autres classes de hauteurs puisque *si3* est omniprésent dans la pièce tant par le jeu de l'instrumentiste, que par le son diffusé en continu⁵⁵. L'autre classe de hauteurs ayant une saillance significativement supérieure est *mi* (0.34). Mais, ceci est dû principalement à un pic lors des sections 4 et 6b (figure 5).

Les analyses qui vont suivre ont été effectuées en divisant la pièce en sept sections. La segmentation, celle proposée par Pierre Michel, est fondée sur les césures ou les suspensions structurellement les plus importantes de façon à « considérer la forme comme une succession de six propositions ou parties (la

⁵¹ Luciano BERIO, *op. cit.*, p. 435.

⁵² Pierre MICHEL, *op. cit.*, p. 269.

⁵³ Le terme classe de hauteurs (*pitch class*) est pris ici dans son acception psychoacoustique (*tone chroma*).

⁵⁴ Richard PARNCUTT, *Harmony: A Psychoacoustical Approach*, Berlin, Springer-Verlag, 1989.

⁵⁵ Un *si3* de faible intensité doit être diffusé pendant toute la durée de la pièce afin de conférer une légère résonance au son du hautbois. La source sonore (un instrument enregistré ou un appareillage électronique) doit être invisible.

sixième pouvant être décomposée en deux) »⁵⁶. Nous considérons donc chaque section comme une gestalt perceptive unifiée par un choix de hauteurs, un type d'écriture, certains modes de jeu, etc. Les données sont moyennées par section et représentées par des histogrammes. L'analyse par section des classes de hauteurs (figure 4, en abscisse les classes de hauteurs, en ordonnée les sections) permet d'obtenir quelques précisions intéressantes sur la saillance de hauteurs difficilement observables avec la seule partition. La classe *si* reste saillante tout au long de la pièce sauf lors des sections 4 (0.82) et 6b (0.92), c'est-à-dire au milieu et à la fin de la pièce, au moment où la classe *mi* tend à s'imposer. Globalement, la saillance perceptive des hauteurs s'organise autour de trois classes de hauteurs *si/do/si* dont la saillance reste relativement constante. Ce sont les trois premières hauteurs du champ harmonique déterminé par Berio. Ensuite, vient un deuxième groupe de classes de hauteurs, *mi /mi/fa*, qui devient saillant à partir de la troisième section. Ce sont les classes qui interviennent dans la succession des hauteurs du plan harmonique au rang 9, 10, 11. Certaines classes de hauteurs n'ont un poids perceptif que ponctuellement. Ainsi en est-il de *ré* dans la troisième section (0.25), de *mi* dans la quatrième (0.67), de *fa* dans la section 6b (0.26) et de *ré* dans la section 3 (0.25). Toutes les classes de hauteurs sont présentes dans la dernière section (S6b), mais assez faiblement, excepté *si* (0.91) et *mi* (0.36).

<i>la</i>	<i>la</i>	<i>si</i>	<i>do</i>	<i>do</i>	<i>ré</i>	<i>mi</i>	<i>mi</i>	<i>fa</i>	<i>fa</i>	<i>sol</i>	<i>sol</i>
0.10	0.14	1.06	0.17	0.08	0.12	0.12	0.34	0.14	0.15	0.14	0.08

Figure 4

⁵⁶ Pierre MICHEL, *op. cit.*, p. 272.

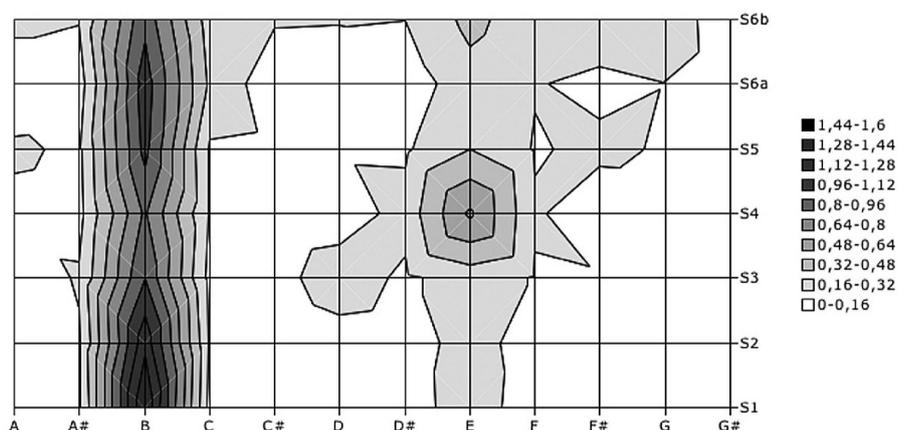


Figure 5 : Saillance des classes de hauteurs par section

Dissonances sensorielle et harmonique

La *Sequenza VII* comporte plusieurs facteurs capables de suggérer des phénomènes de tension/détente tels que la densité chromatique, les intervalles, le choix du registre ou le recours à certains modes de jeu (sons multiphoniques, sons forcés, trémolos, doubles trilles). Nous pouvons donc nous interroger sur la façon dont ces phénomènes se traduisent dans la pièce et sur les facteurs qui les déterminent. Depuis la *Théorie physiologique de la musique* d'Hermann von Helmholtz⁵⁷, on considère que la notion de consonance est fondée sur l'interaction de deux composantes perceptives : la consonance sensorielle (*Klangverwandtschaft*) et la consonance « harmonique » (*Konsonanz*). La première est liée aux sensations auditives, alors que la seconde dépend de facteurs plus perceptifs et cognitifs. Ernst Terhardt attribue la consonance sensorielle « à la perception de fluctuations du son (rugosité) et à la clarté spectrale » et la consonance harmonique « à la perception de hauteurs de signaux complexes »⁵⁸. La notion de rugosité (*roughness*) a donné lieu à de nombreux travaux en psychoacoustique⁵⁹ montrant qu'elle repose sur la

⁵⁷ Hermann von HELMHOLTZ, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig, 1863. Fac-similé de l'édition française (1868) *Théorie physiologique de la musique fondée sur l'étude des sensations auditives*, traduit de l'allemand par M. G. Guérault, Paris, Editions Jacques Gabay, 1990.

⁵⁸ Ernst TERHARDT, « The concept of Musical Consonance: A Link between Music and Psychoacoustics », *Music Perception* 1/3 (1984), p. 292.

⁵⁹ Reiner PLOMP et Willem J. M. LEVELT, « Tonal consonance and critical bandwidth », *Journal of the Acoustical Society of America* 38 (1965), p. 548-560. Aklo KAMEOKA et Mamoru KURIYAGAWA, « Consonance Theory, part I, II », *Journal of the Acoustical Society of America* 45 (1969),

différence de fréquence relativement à une largeur de bande critique. Daniel Pressnitzer⁶⁰ et ses collègues ont montré que des stimuli orchestraux complexes (tirés de la pièce *Streamlines* de Joshua Fineberg), dépourvus de fonctions harmoniques au sens tonal du terme, peuvent évoquer à des sujets musiciens et non musiciens des sensations de tension et de détente en fonction de leur degré de rugosité. La clarté spectrale (*spectral sharpness*) est relative à la présence d'une grande quantité d'énergie dans le haut du spectre qui peut rendre le timbre dur, acide, « pointu ». La consonance harmonique est peut-être plus difficile à mesurer avec des descripteurs psychoacoustiques. Mais, il est possible d'obtenir des indices avec des descripteurs comme la clarté tonale (*tonalness*) et la multiplicité (*multiplicity*). La clarté tonale est définie par Richard Parncutt (1989) comme « la perceptibilité de la composante la plus audible d'un son »⁶¹. Elle reflète l'audibilité de la fondamentale virtuelle d'un accord (un accord parfait aura une plus grande clarté tonale qu'un accord de septième diminuée). La multiplicité (*multiplicity*) mesure selon Parncutt « le nombre de composantes perçues simultanément dans les présentations courtes et isolées de sons musicaux »⁶². Elle correspond au nombre de composantes perceptibles dans une polyphonie verticale ou horizontale et reflète la complexité de la texture et de l'harmonie. Nous allons employer ces quatre descripteurs psychoacoustiques : rugosité, clarté spectrale, clarté tonale et multiplicité en tant qu'attributs perceptifs mesurables permettant d'étudier les phénomènes de tension/détente dans les *Sequenza VII*. La méthode consiste donc à procéder aux mesures avec un logiciel (ces descripteurs psychoacoustiques sont implémentés dans le logiciel Psysound 3⁶³), puis à interpréter les résultats en fonction des éléments d'écriture présents dans la partition.

La figure 6 présente le taux moyen de rugosité pour chaque section. On peut observer un accroissement de la rugosité jusqu'à la section 4, une chute lors de la section 5, un taux équivalent à celui de la section 3 lors de la section 6a et enfin une nouvelle chute, plus importante, pendant la dernière section. Le profil global de rugosité est cohérent avec le déploiement des hauteurs (figure 3). Berio introduit onze des douze classes de hauteurs entre les sections 1 et 4. Il y a donc un accroissement de la densité harmonique qui se traduit par une augmentation de la rugosité. La section 5, sorte de reprise variée des sections 2 et 3, contenant de nombreux sons répétés ou tenus et relativement peu de traits balayant l'espace chromatique voit son taux de rugosité baisser. La forte chute de la rugosité lors de la dernière section (6b) peut s'expliquer par un ralentissement de l'activité

p. 1451-1469. William HUTCHINSON et Leon KNOPOFF, « The acoustic components of western consonance », *Interface 7* (1978), p. 1-29.

⁶⁰ Daniel PRESSNITZER, Stephen McADAMS, Suzanne WINSBERG et Joshua G. FINEBERG, « Tension et rugosité dans une musique non tonale », *JIM: Journées d'informatique musicale* (1996), <http://mediatheque.ircam.fr/articles/textes/Pressnitzer96a/>.

⁶¹ Richard PARNCUTT, *op. cit.*, p. 59.

⁶² Richard PARNCUTT, *op. cit.*, p. 100.

⁶³ Psysound 3 est développé en Matlab par Densil CABRERA et Sam FERGUSON, <http://psysound.wikidot.com/html> (consulté le 19 mars 2009).

(suspensions sur des notes tenues), mais peut-être également par la présence de l'intervalle de quinte dans les sons multiphoniques. En tout cas, la mesure de la rugosité fait apparaître une baisse de la dissonance sensorielle lors de cette dernière section qui se traduit par une fonction de détente.

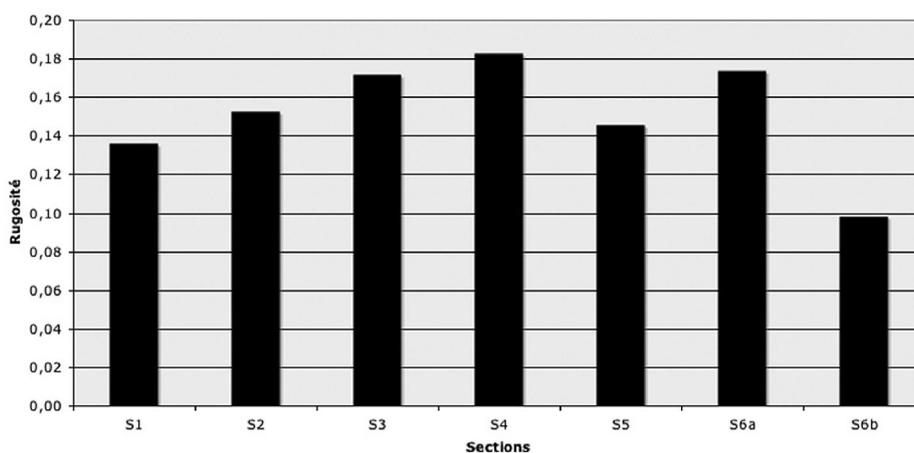


Figure 6 : Rugosité (*roughness*) par section

Les mesures de clarté spectrale sont reportées à la figure 7. Le profil global est sensiblement différent de celui de la rugosité. Les sections 5 et 6a se détachent par leur degré plus important de clarté spectrale alors qu'inversement les sections 4 et 6b se démarquent par leur faible taux de clarté spectrale. Les sections 5 et 6a sont celles où le registre aigu est systématiquement exploité, au détriment du registre grave. La section 6a représente selon Berio le point culminant de la pièce par l'apparition du douzième son dans le registre suraigu (*sol5*). La répétition de cette note, parfois agrémentée de trémolos, produit un effet de tension dû tant à la sonorité qu'à la difficulté d'émission requise. Inversement, une partie de la section 3 et la section 4 privilégient le registre grave (*ré3, mi 3, mi3*). La chute de la clarté spectrale lors dernière section provient moins du registre que de l'intensité faible (nuances autour de *p* ou *mf*) qui implique un timbre de hautbois contenant moins harmoniques aigus. La détente lors de la dernière section, constatée avec la mesure de rugosité, semble être renforcée par l'adoucissement de la sonorité.

La *Sequenza VII* ne comporte pas, *stricto sensu*, d'accords puisqu'il s'agit d'une pièce pour instrument monodique. Cependant, l'idée de polyphonie latente est au cœur de cette *Sequenza*. La polyphonie est réalisée par plusieurs artifices : premièrement la diffusion à faible volume d'un *si3* pendant toute la durée de la pièce afin de conférer une légère résonance au son du hautbois, deuxièmement, le recours à des modes de jeu qui produisent des effets harmoniques (sons multiphoniques, sons forcés, doubles trilles), et troisièmement la succession rapide de hauteurs que la mémoire à court terme tend à regrouper en harmonies sous-jacentes. Il est donc intéressant de mesurer combien le son pivot (*si3*), que l'on peut

considérer comme une fondamentale virtuelle, est perceptivement présent. La clarté tonale (*complex tonalness*) correspond à l'audibilité de la fondamentale virtuelle d'un accord ou d'un agrégat. La figure 8 présente le degré de clarté tonale moyen pour chaque section. On observe une décroissance globale de la clarté tonale qui reflète l'affaiblissement progressif de la fondamentale virtuelle. La première section, ne comportant pratiquement que des variations de timbre et de rythme sur la note pivot (*si3*), possède logiquement le taux le plus élevé de clarté tonale. Inversement, les deux dernières sections (6a et 6b), dans lesquelles les changements de hauteurs sont quasiment constants, ont les taux les plus faibles de clarté tonale. La saturation chromatique perturbe la perception de la fondamentale virtuelle, malgré la présence sous-jacente du son pivot toujours diffusé en coulisse.

Le descripteur de multiplicité apporte des indications sur la façon dont la polyphonie virtuelle est perçue en tant que flux harmonique produit par les modes de jeu. Logiquement, l degré de multiplicité le plus élevé se situe lors de la dernière section (6b) puisque c'est celle qui comporte le plus de multiphoniques et de doubles trilles (10 sur un total de 14). L'élévation du taux de multiplicité lors des sections 3, 4, 5 et 6a, moins flagrante que celle de la dernière section, semble plutôt reposer sur l'abondance des sauts véloces de hauteurs que l'oreille tend à fusionner en harmonies verticales. Sans surprise, les deux premières sections ont les taux les plus faibles car le jeu de l'instrumentiste se limite aux cinq premières classes de hauteurs avec de nombreuses répétitions du son pivot (*si3*).

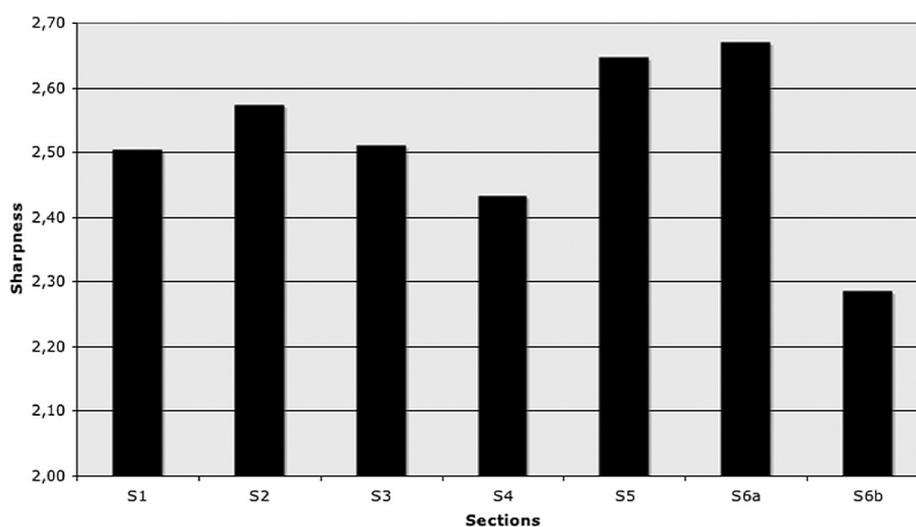


Figure 7 : Clarté spectrale (*spectral sharpness*) par section

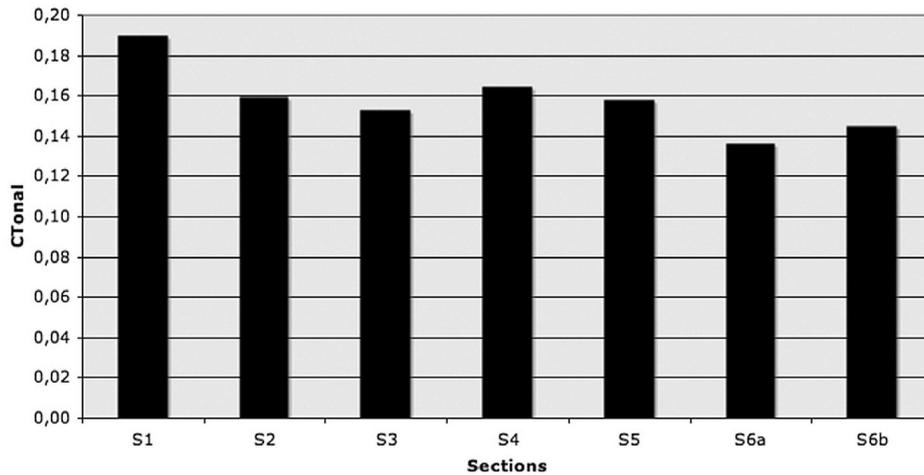


Figure 8 : Clarté tonale (*complex tonalness*) par section

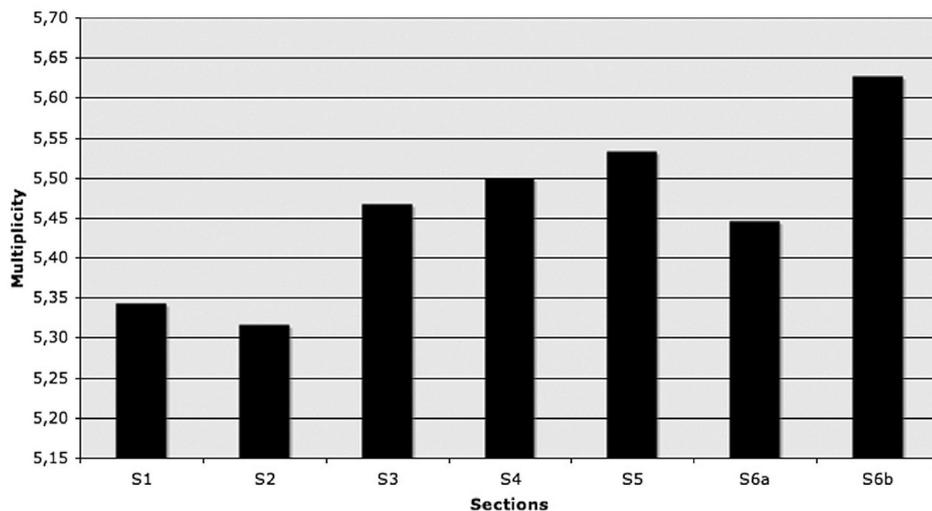


Figure 9 : Multiplicité (*multiplicity*) par section

Il est difficile de rendre compte de la combinaison complexe des dissonances sensorielle et harmonique à un niveau local. On peut cependant tirer quelques remarques des résultats obtenus sur un plan global. Lors des premières sections les deux types de dissonance convergent dans la direction d'un accroissement. Les sections centrales (S4 à S6a) sont celles où la tension est à son maximum les deux types de dissonance se renforçant l'une l'autre. La dernière section (6b) résout la tension accumulée tant par le statisme dû à la verticalisation (points d'orgue avec des sons multiphoniques) que par la chute de la rugosité et de la clarté spectrale. Le

schème tensionnel de la *Sequenza VII* s'apparente donc à une enveloppe archétypale montée, sommet, chute bien connue de l'auditeur occidental.

Conclusion

Le paradigme sub-symbolique offre une nouvelle voie pour l'analyse musicale assistée par ordinateur jusqu'alors centrée sur le traitement computationnel de données symboliques ou sur l'analyse sonographique. Elle s'appuie, non sur le signal acoustique lui-même, mais sur les données modélisant la réponse du système auditif à ce signal. Contrairement au paradigme symbolique, le paradigme sub-symbolique constitue une approche ascendante (*bottom-up*) qui de la sensation auditive s'achemine vers la signification. L'extraction d'informations sub-symboliques opérée par les descripteurs psychoacoustiques offre la possibilité d'approcher la perception de plusieurs dimensions musicales telles que le temps (tempo, attaques, durées), la hauteur (rugosité, saillance, multiplicité, clarté tonale, etc.) et le timbre (brillance, inharmonicité, enveloppe, clarté spectrale, etc.). Ces outils informatiques semblent particulièrement pertinents pour l'analyse des processus à grande échelle. Ainsi, l'analyse de la *Sequenza VII* de Berio a permis de répondre à des questions précises sur la répartition des attaques en fonction du canevas temporel de la pièce, sur la saillance perceptive de chaque classe de hauteurs et sur les phénomènes de tension/détente produits par la combinaison des dissonances sensorielle et harmonique. Ce travail, se situant dans une optique exploratoire, ne constitue qu'une approche de l'utilisation des descripteurs psychoacoustiques pour l'analyse musicale. De nombreuses questions restent encore en suspens concernant le choix et la pertinence des descripteurs en fonction des contextes musicaux, la présentation des résultats et l'intégration de ces outils dans les méthodes actuelles d'analyse musicale.