



HAL
open science

Une proposition pour l'analyse des musiques électroacoustiques de Xenakis à partir de l'utilisation de descripteurs audio

Mikhail Malt

► **To cite this version:**

Mikhail Malt. Une proposition pour l'analyse des musiques électroacoustiques de Xenakis à partir de l'utilisation de descripteurs audio. Makis Solomos Iannis Xenakis, La musique électroacoustique, L'Harmattan, pp.159-198, 2015, 978-2-343-06696-7. hal-01265273

HAL Id: hal-01265273

<https://hal.science/hal-01265273v1>

Submitted on 1 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mikhail Malt, « Une proposition pour l'analyse des musiques électroacoustiques de Xenakis à partir de l'utilisation de descripteurs audio », Makis Solomos (ed.) *Iannis Xenakis, La musique électroacoustique*, L'Harmattan, 2015, p. 159-198.

Une proposition pour l'analyse des musiques électroacoustiques de Xenakis à partir de l'utilisation de descripteurs audio

Mikhail Mal (Ircam – STMS, IReMus – Mint– Paris Sorbonne, France)

ABSTRACT

Using low level audio descriptors, we propose a methodology for musical analysis, using four Xenakis' electroacoustic works composed at GRM (Musical Research Group). In this methodology we propose a model of the concept of mass sound by a graphic representation, BSTd (Brightness-Standard deviation) (Malt, Jourdan 2009).

1. INTRODUCTION

Cette communication constitue la première partie d'une recherche plus étendue sur l'utilisation et l'évaluation des descripteurs audio et de méthodes issues des recherches en MIR (Musical Information Retrieval) en analyse musicale. Nous la concevons comme un exercice de style analytique, nous basant sur l'utilisation de la représentation graphique de l'évolution d'un certain nombre de descripteurs audio en analyse musicale.

Il est important de signaler que l'utilisation de ces représentations est pensée comme une « aide », dans une première phase de description du phénomène étudié. Les descripteurs audio étant, par définition, une représentation de paramètres audio, ils sont un point de vue partiel, leur fonction étant d'apporter un certain nombre d'informations pour le processus d'analyse musicale, en allant dès les phases de segmentation à la phase d'interprétation. L'utilisation de l'information apportée par chacune de ces représentations sera conditionnée et devra être cautionnée, avalisée ou réfutée par d'autres éléments analytiques.

L'hypothèse principale de cette étude en particulier est de considérer qu'un concept implicite, ou explicite, dans le parcours compositionnel de Xenakis est le concept de « masse sonore », comme un paramètre d'une dénomination plus vaste que Xenakis utilise, soit le terme d'« entité sonore », dénomination qui lui sert à désigner un « objet sonore » dans son ensemble : « La musique peut donc être définie comme une organisation de ces opérations et de ces relations élémentaires entre des êtres ou entre des fonctions d'êtres sonores » (Xenakis, 1963, 16).

Nous proposerons une représentation graphique (fondée sur l'utilisation de descripteurs audio), la BSTd, susceptible de prendre en compte, ou de représenter, ne serait-ce que partiellement cette dimension des entités sonores, que nous nommerons la « masse sonore ». Cette représentation et ce concept seront explicités dans la suite de ce texte.

Il est important de signaler que ce point de vue de la masse sonore, comme nous l'entendons, est d'un niveau hiérarchique plus bas que, par exemple, celui des « complexes sonores », tels que définis par M. Solomos (Solomos 1994). La « masse sonore » sera vue comme une morphologie physique, plutôt qu'une interprétation musicale, dont l'analyse donnera la signification et la fonction musicale en fonction du contexte. Ce que nous proposons, en première instance, est une représentation, un outil de description pour des phases préliminaires de l'analyse.

2. METHODOLOGIE

2.1. Le corpus d'œuvres

Nous nous concentrerons sur les quatre œuvres pour bande seule de la phase « concrète » de Xenakis [Harley 2002, 34], *Diamorphoses* (1957), *Concret PH* (1958), la version concert d'*Orient-Occident* (1960) et *Bohor* (1962). Ce choix a été fait pour nous concentrer sur une phase de travail du compositeur, présentant une unité de techniques de manipulation. L'aspect « musique concrète », avec un vocabulaire restreint de manipulations, indique une manière de faire, un artisanat électroacoustique, fondée sur la manipulation basique de matériau sonore avec un vocabulaire restreint d'opérations : la découpe, le collage, la transposition, etc. Cet aspect nous permettra de mieux comprendre comment l'œuvre finale prend sa forme et de comment les différents paramètres inférés procèdent plus d'une volonté et d'un métier du compositeur que des traitements audio.

2.2. Canal ou multicanal ?

Pour cette première étude, nous ne prendrons pas en compte l'aspect spatial de la diffusion des œuvres, nous travaillerons sur un fichier audio monophonique, résultant de l'aplatissement des canaux originaux. Nous sommes conscients que cette décision sera à l'origine d'un problème méthodologique qui aura pour conséquence le « moyennage » de certains résultats, et/ou de certains aspects de l'écriture qui se développent spatialement. Cela sera très sensible pour *Bohor* (8 pistes) et même pour la version stéréo de *Concret PH*, où par exemple il est possible d'identifier des imitations de rythmes entre les canaux gauche et droit, sans doute issus de copies d'éléments de la bande à des temporalités différentes. Cependant, comme nous le verrons, les fichiers audio aplatis nous retourneront un certain nombre d'informations pertinentes.

Nous espérons pouvoir régler ce problème méthodologique dans des travaux ultérieurs, où chaque canal sera l'objet d'une analyse indépendante¹.

2.4. Les types de fichiers audio

Nous avons travaillé principalement avec des fichiers audio numériques non compressés au taux d'échantillonnage de 44 100 Hz et à la précision de 16 bits. Il n'y a pas à notre connaissance (peut être une faille dans notre recherche bibliographique) d'études sur l'influence de la compression sur le résultat d'analyses des descripteurs audio². Par contre, il est facile de se rendre compte que l'utilisation de fichiers compressés (compression destructive : mp3, mp4, etc.) pour des analyses spectrales comporte des désavantages par le simple fait que la compression est visible (dans les analyses spectrales) comme un ensemble de régions vides. Les sonagrammes ressemblent à de vrais « gruyères »... Les Figure 1 et Figure 2 présentent des analyses spectrales du même extrait de « Symphonie pour un homme seul » de Pierre Henry, mouvement 12 – « Strette » de 15.4 s à 16.3 s (de 0 à 10 000 Hz) [Shaeffer 1990, 2-12]. Nous pouvons remarquer dans la Figure 2 les « trous » présents entre 5000 et 7500 Hz.

¹ Nous aimerions citer le fait qu'au moment de l'écriture de cet article (2014), Pierre Couprie faisait des analyses en étudiant l'évolution de descripteurs audio, notamment le centroïde spectral sur les différents canaux pour vérifier la corrélation entre les différentes pistes audio.

² Jusqu'à la date de l'écriture de cet article.

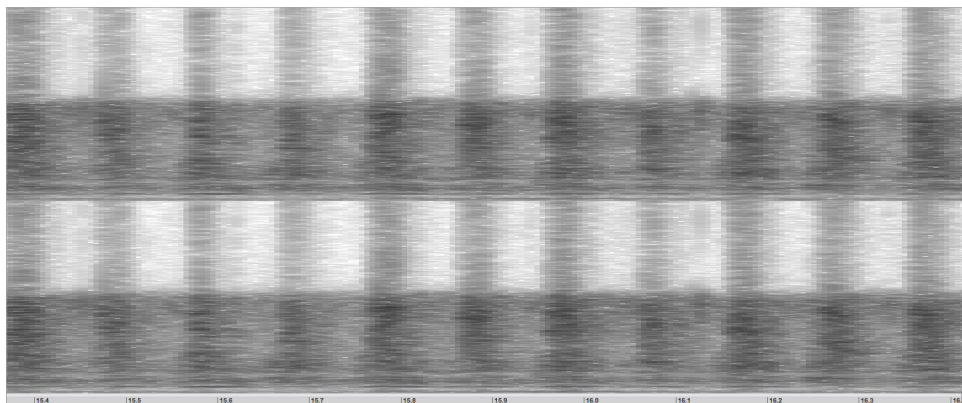


Figure 1. Spectrogramme du fichier original³.

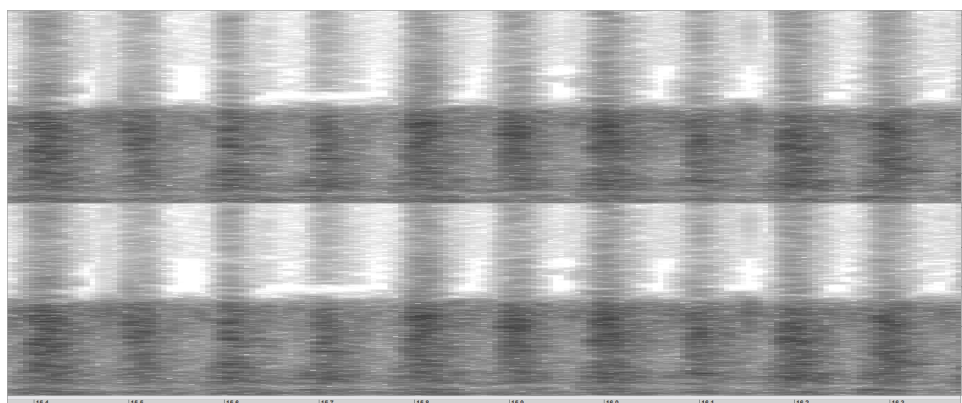


Figure 2. Spectrogramme du fichier audio compressé⁴.

2.5. L'environnement de travail

Comme il n'existe pas d'outil informatique, ou d'environnement d'analyse spécifique pour traiter ce type de recherche⁵, nous avons développé un outil de travail dédié, utilisant principalement l'environnement de programmation graphique *Max*⁶, la librairie *zsa.descriptors*⁷ de calcul de descripteurs, l'objet *IrcamDescriptors*⁸, un ensemble de fonctions « ad hoc » pour le calcul de rugosité (écrites en JavaScript) et l'objet *MUBU*⁹. Pour la représentation en forme d'onde et l'analyse de sonagrammes, nous avons utilisé le logiciel *Audiosculpt*¹⁰. Pour la présentation de données le logiciel *EAnalysis*¹¹ de Pierre Couprie, et pour des calculs statistiques *R*¹².

2.6. Les descripteurs Audio

Un descripteur audio est un paramètre, uni ou multidimensionnel, ramenant une dimension particulière du signal audio, à un (ou plusieurs) paramètre(s) numérique(s). Son évolution

³ Fichier audio aiff extrait de Shaeffer (1990, 2-12), Taux d'échantillonnage 44100 Hz, précision 16 bits. Spectrogramme (transformée de Fourier rapide) réalisé avec le logiciel Audiosculpt (© Ircam), taille de fenêtre 2048 échantillons, pas d'analyse 256 points, fenêtre d'analyse « blackman ».

⁴ Compression réalisée avec le logiciel Audacity (<http://audacity.sourceforge.net/>) utilisant la librairie « libmp3lame.dylib » (<http://lame.sourceforge.net/>), débit de 128 kbits/sec.

⁵ Notons quand même l'existence de Sonic Visualizer (www.sonivisualiser.org) avec son ensemble de fonctionnalités permettant d'extraire de nombreux descripteurs. Cependant, il n'est pas possible dans ce logiciel de représenter la BStD.

⁶ <http://cyclimg74.com/>

⁷ http://www.e--j.com/?page_id=83. Voir aussi [Malt & Jourdan 2008] et [Malt & Jourdan 2009]

⁸ © Ircam, objet externe Max développé par Geoffroy Peeters et Diemo Schwartz)

⁹ <http://imtr.ircam.fr/imtr/MuBu>. Voir [Schnell & al. 2009].

¹⁰ © Ircam, <http://anasyth.ircam.fr/home/software/audiosculpt>.

¹¹ http://logiciels.pierrecouprie.fr/?page_id=402.

¹² <http://www.r-project.org/>

temporelle peut être utilisée pour représenter, caractériser, illustrer, etc. un flux ou une évolution d'évènements sonores. Initialement prévus comme des moyens pour identifier des caractéristiques sonores et ensuite pour faciliter l'extraction d'information dans de vastes bases de données de musique enregistrée, ils sont aujourd'hui suffisamment développés, et surtout accessibles, pour que soit considérée leur intégration dans des stratégies analytiques prenant en compte les textes musicaux et le signal audio provenant d'enregistrements.

2.7. Analyses

Pour chaque œuvre nous avons calculé un ensemble de quarante-quatre (44) descripteurs, plus la construction d'une représentation composée de « masse sonore ». Les paramètres de l'analyse étaient les mêmes pour toutes les analyses : Taux d'échantillonnage 44 100 Hz, taille de fenêtre 2048, pas d'avancement 512, fenêtre « blackman ». Les descripteurs utilisés étaient :

- 1) Trente-sept IrcamDescriptors [Peeters 2003] : HarmonicSpectralRolloff, HarmonicSpectralSlope, HarmonicSpectralDecrease, HarmonicSpectralVariation, HarmonicSpectralKurtosis, HarmonicSpectralSkewness, HarmonicSpectralSpread, HarmonicSpectralCentroid, HarmonicSpectralDeviation, HarmonicOddToEvenRatio, Inharmonicity, Noisiness, NoiseEnergy, HarmonicEnergy, PerceptualSpectralRolloff, PerceptualSpectralSlope, PerceptualSpectralDecrease, PerceptualSpectralVariation, PerceptualSpectralKurtosis, PerceptualSpectralSkewness, PerceptualSpectralSpread, PerceptualSpectralCentroid, PerceptualOddToEvenRatio, PerceptualSpectralDeviation, Sharpness, Spread, Loudness, SpectralRolloff, SpectralSlope, SpectralDecrease, SpectralVariation, SpectralKurtosis, SpectralSkewness, SpectralSpread, SpectralCentroid, TotalEnergy et SignalZeroCrossingRate.
- 2) Cinq Zsa.descriptors ([Malt & Jourdan 2008] et [Malt & Jourdan 2009]) : Spectralcentroid, SpectralSpread, Energy, SpectralFlux et SpectralFlatness.
Dans le cadre du calcul du « SpectralFlatness » avec la librairie zsa.descriptors, nous avons utilisé une seule bande de fréquence de 20 à 20000Hz.
- 3) Deux fonctions *javascript*, implémentation de deux fonctions (RoughnessF et RoughnessV) pour le calcul de la rugosité, à partir des travaux de Daniel Prestnitzer [Prestnitzer & al. 2000] et de Pantelis N. Vassilakis [Vassilakis 2001].

2.8. Evaluation

En ce qui concerne l'évaluation des résultats de notre recherche et du choix des descripteurs, nous avons travaillé sur trois points :

- 1) La vérification de l'alignement des données issues de l'analyse par des descripteurs, ou de la représentation de « masse sonore » avec des données issues d'analyses musicologiques de la littérature et de la perception des évènements.
- 2) La recherche de directions ou trajectoires dans l'évolution des descripteurs et/ou des « masses sonores ».
- 3) Finalement, la possibilité d'extraire un diagramme ou schéma formel à partir d'une segmentation induite par ces représentations.

Dans le cadre de cette communication, nous présenterons une méthodologie plutôt qualitative fondée sur des analogies et des similarités visuelles et auditives, destinées à l'utilisation des musicologues. Cependant, nous sommes en train de développer une méthodologie quantitative, fondée sur l'analyse statistique des descripteurs, qui devra être communiquée prochainement.

2.9. Les « masses sonores »

Nous définirons une masse sonore (en nous inspirant des concepts de Shaeffer et de Smalley) comme un *registre spectral* et une *texture*. Le registre spectral sera défini par le couple $\left[\left(C + \frac{S}{2} \right), \left(C - \frac{S}{2} \right) \right]$, où C est la fréquence du « centroïde spectral », S l'écart-type spectral, $\left(C + \frac{S}{2} \right)$ étant la borne supérieure du registre et $\left(C - \frac{S}{2} \right)$ la borne inférieure. La texture sera représentée par une ou plusieurs caractéristiques sonores, telles que l'amplitude du signal, son énergie, le flux spectral (ou variation spectrale), la « spectral flatness », la rugosité, ou autres. L'évolution de la « masse sonore » sera représentée par une forme graphique (Figure 3), où le registre spectral sera affiché sur l'axe vertical (les ordonnées), l'axe des abscisses représentera le temps et l'axe z (par la couleur) représentera la texture du registre spectral à un temps donné. Dans un travail antérieur [Malt & Jourdan 2009] nous avons nommé cette représentation BStD (brightness-standard deviation)¹³. Il est important de signaler que la forme qui est représentée n'est pas un sonagramme, mais une forme qui par la position de son barycentre indique la brillance, et par son étendue verticale le « volume spectral »¹⁴.

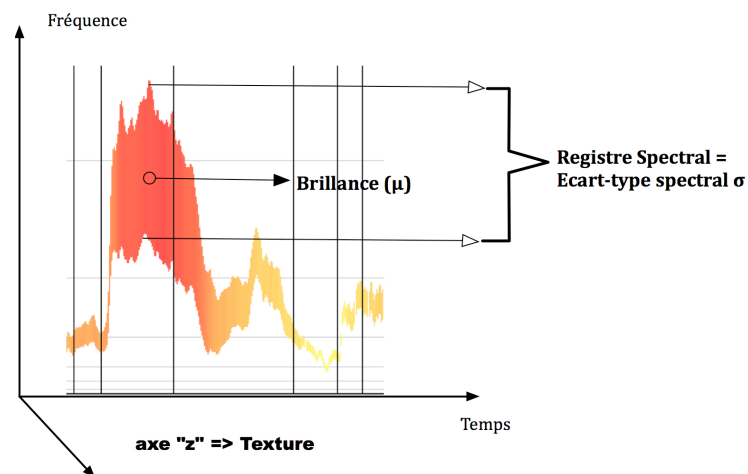


Figure 3. La représentation de l'évolution de la « masse sonore », avec la BStD.

Le but principal de cette définition, et de cette représentation, est de proposer un outil pour l'étude du timbre dans le cadre de l'analyse musicologique.

3. DIAMORPHOSES (1957)

Diamorphoses est la première pièce électroacoustique de Xenakis. L'œuvre originale est une bande sur quatre canaux d'une durée de 7 minutes. Il existe plusieurs analyses de cette œuvre. Nous nous fonderons sur deux travaux, celui de Harley [Harley 2002] et celui de Valsamakis [Valsamakis 2009].

Commençons avec une représentation de l'évolution de la masse sonore, par la BStD (Figure 4), où la texture (la couleur) soit la troisième dimension est représentée par l'énergie du signal audio.

¹³ Voir aussi [Siedenburger 2009].

¹⁴ Pour les rapports entre l'écart type spectral et la notion de « volume sonore » voir [Chiasson 2007].

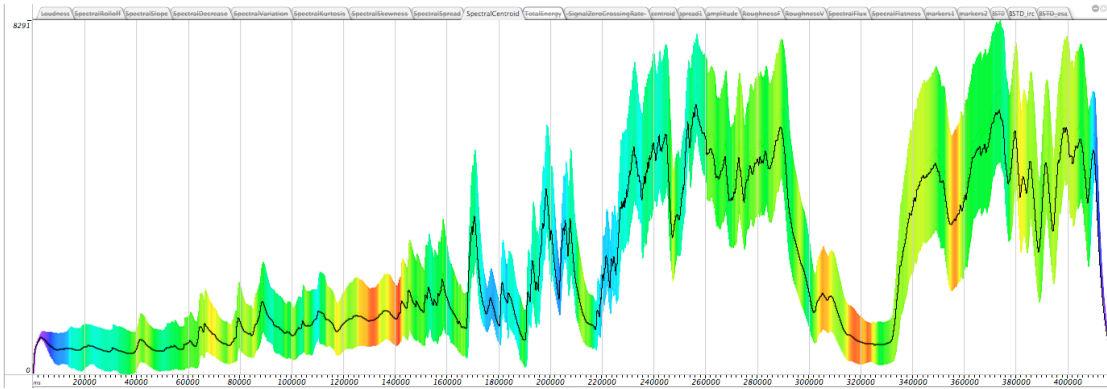


Figure 4. L'évolution de la « masse sonore » de *Diamorphoses* (1957).

Premier point, comment interpréter cette représentation ?

En suivant l'évolution de la forme générée, nous pouvons remarquer 3 grandes catégories morphologiques résumées dans le Tableau 1 et présentées dans la Figure 5.

Catégories	Caractéristiques
1	Trajectoire ascendante ou descendante de la forme avec une évolution moindre du registre spectral, approximée par une droite
2	Large oscillation du centroïde spectral (+/- 4000 Hz), avec un registre spectral mince (+/- 1200 Hz).
3	Registre spectral épais (+/- 1000-1200 Hz), avec une tendance à une trajectoire horizontale, ou plutôt stable.

Tableau 1. Catégories morphologiques basiques, d'après l'évolution de la « masse sonore ».

Un point important à signaler est le fait que les morphologies que nous sommes en train d'identifier ne correspondent pas forcément à une fonction musicale pour le moment, et comme nous l'avons déjà mentionné, que la « fonction musicale » d'une morphologie sera contextuelle, et définie à une étape postérieure de l'analyse.

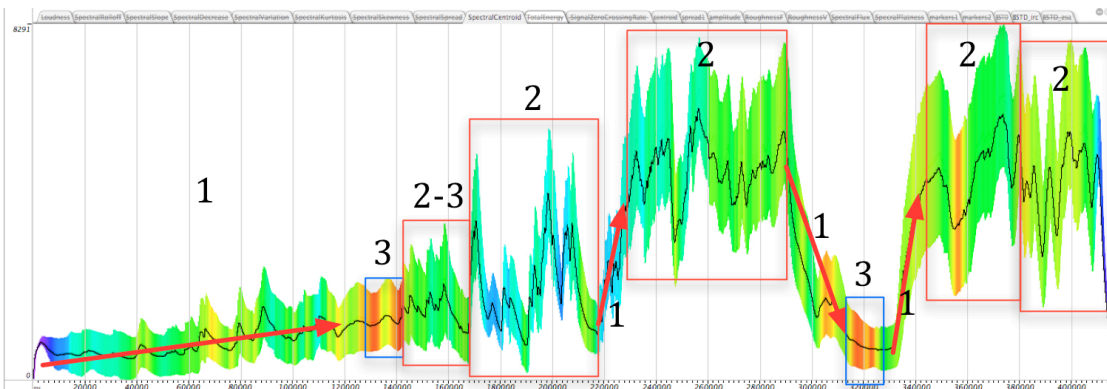


Figure 5. Les trois morphologies (*Diamorphoses*).

Ces trois morphologies s'organisent de la manière suivante :

1 3 2/3 2 1' 2' 1'' 3' 1''' 2'' 2'''', où 2/3 indique une morphologie hybride entre les catégories 2 et 3, soit une oscillation avec un registre spectral épais.

Dans la Figure 6, nous pouvons remarquer comment les segmentations proposées par les morphologies repérées concordent avec des segmentations de la littérature [Harley 2002, 35] et en apportent quelques précisions. Par exemple, la séquence entre 130.57 s et 142.41 s (la première morphologie « 3 »), propose une morphologie de type « bloc », avec un registre

spectral plutôt stable, nous indiquant une fonction musicale de « pont » ou « transition » entre la première et la deuxième partie de cette œuvre. La même catégorie morphologique apparaît entre 315.05 s et 323.47 s (la dernière morphologie « 3' »), montrant la même fonction musicale et structurale de pont, d'élément de transition se situant, aussi comme l'élément précédent, juste avant le début de la dernière section.

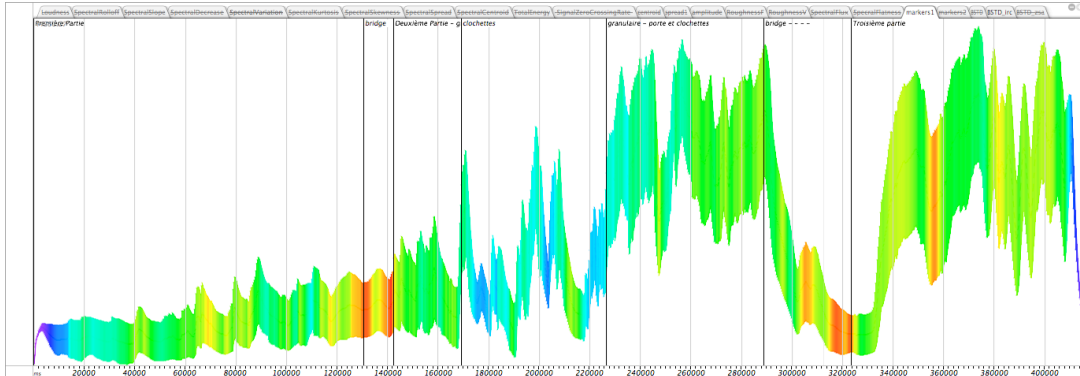


Figure 6. Evolution de la « masse sonore » avec des segmentations.

À la lumière de ces informations, et sans rentrer dans le détail des matériaux utilisés, nous pouvons inférer une structure tripartite (comme cela est proposé par la littérature), avec une première partie caractérisée par une lente progression du centre du registre spectral (progression du centroïde et de l'écart-type spectral, catégorie, et de sons soutenus, avec une transition [pont] pour la prochaine section [catégorie 3]. Une deuxième partie [marquée par l'entrée d'un bruit de claquement de porte, et la morphologie 2/3] caractérisée par un mouvement ascendant [et oscillant] de la masse sonore jusqu'à un point culminant entre 3 min 46 s et 4 min 49 s et une section de transition [entre 4 min 49 s et 5 min 23 s] divisée en deux parties. Cette deuxième partie contraste avec la première par la discontinuité des matériaux. Comme le remarque Harley [Harley 2002, 35] *Diamorphoses* vient du grec, signifiant continuité/discontinuité, soit les deux aspects de l'être.

Une dernière section, divisée en deux parties. La première partie de la troisième section entre 5 min 23 s et 6 min 17 s [approximativement], caractérisée par un accroissement et une trajectoire ascendante de la masse sonore, notamment du registre spectral [Figure 7] ; et la deuxième partie de la troisième section, caractérisée par une oscillation de la masse sonore et d'une diminution de l'étendue du registre spectral. Cette dernière partie, de la troisième partie, ne semble pas avoir de trajectoire globale, mais plutôt d'osciller autour de 4000-4200 Hz, renvoyant à une sensation d'immobilité.

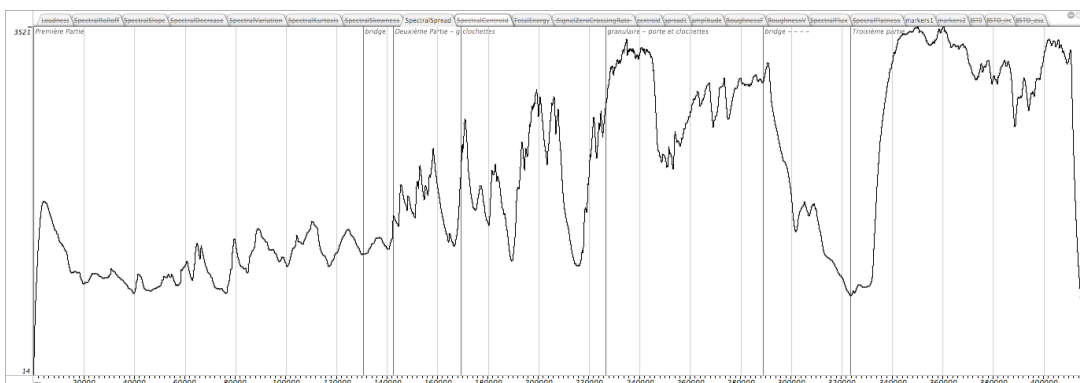


Figure 7. Evolution de l'écart-type spectral.

Il s'avère utile de compléter nos informations par d'autres descripteurs. Dans ce cas précis, la variation perceptive de flux spectral¹⁵ (Figure 8).

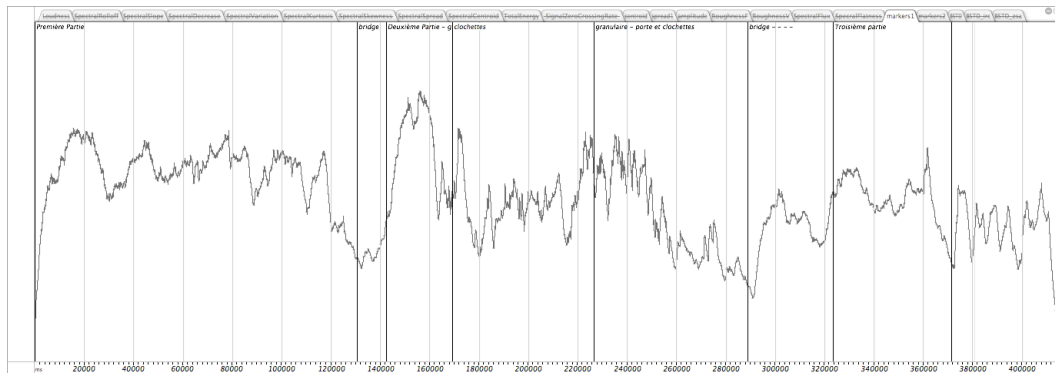


Figure 8. Evolution de la « Perceptual Spectral Variation ».

Dans la première partie, nous remarquons une oscillation du flux spectral, avec une légère tendance à une trajectoire descendante. Cela indique une masse sonore qu'évolue avec une densité de mouvement interne plutôt constante, et une décroissance dans cette densité interne à la fin. Avant la deuxième partie, nous remarquons une baisse du flux spectral, nous renvoyant à un flux audio avec une tendance à l'immobilité, c'est le premier pont ou transition entre la première et deuxième section. Le plus intéressant est la deuxième partie de la troisième section (Figure 9). Nous remarquons aussi une oscillation plus marquée du flux spectral. Par une oscillation continue, sans trajectoire privilégiée, cette dernière partie fonctionne comme une coda, une cadence conclusive. Nous nommerons cette partie de « pédale statistiquement statique ». Comme nous le verrons par la suite, cette stratégie sera utilisée plusieurs fois par Xenakis, comme une manière d'arrêter le mouvement, par une redondance de l'information. Le Tableau 2 présente une proposition de description formelle de *Concret PH*.

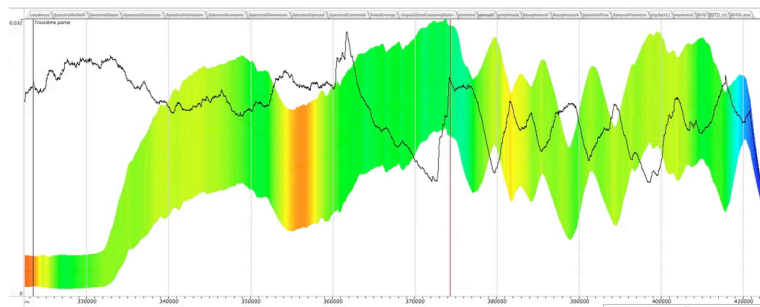


Figure 9. La troisième partie de *Diamorphoses* (masse sonore, courbe = Flux spectral perceptif).

¹⁵ La « Spectral Variation », soit la variation de flux spectral, nous renvoie une information concernant la corrélation qu'existe entre deux analyses spectrales conjointes. C'est-à-dire, si un flux sonore est immobile, cette valeur tend à être basse, puisque les analyses renvoient, à peu près les mêmes valeurs. Par contre pour un flux sonore avec du mouvement interne (de la synthèse granulaire, par exemple) cette valeur tend à s'accroître. La « Perceptual Spectral Variation » est le calcul de la « Spectral Variation », après que le flux audio soit filtré par un modèle de l'oreille moyenne et converti en bandes de Mel ou Bark (voir [Peeters 2003, 4]).

Section I		Section II			Section III	
1	3	2/3 2 1' 2'		1'' 3'	1'''	2'' 2'''
					Introduction - pont	
Continuité avec évolution		Discontinuité – mouvement – Registre spectral moins épais que la première partie			Coda caractérisée par une oscillation des matériaux sans trajectoire définie. Sensation d'immobilité	
Lente évolution de la masse sonore, avec augmentation de la tension due à la montée dans le registre spectral	Pont- la tension s'arrête	2/3 2	p o n t	2' Point culminant	pont	La « pédale » statistiquement statique

Tableau 2. Proposition de description formelle pour *Diamorphoses*.

4. CONCRET PH (1958)

La genèse et l'histoire de la composition de *Concret PH* sont assez bien connues, pour qu'on puisse s'en passer. Juste pour rappel, cette pièce est une des premières utilisant une conception granulaire du son dans sa composition. Xenakis, se fondant sur les travaux de Denis Gabor, reconstruit une texture sonore à partir d'une multitude de très courts échantillons sonores. Créée en 1958 en version mono, elle a été révisée en 1961 pour en créer une version stéréo et remixée en 1969 pour en faire une version quadraphonique [Harley 2002, 35-36]. Selon Harley [Harley 2002, 37], Xenakis a produit de multiples versions de toutes ses pièces pour bande, jusqu'à la *Légende d'Eer*, pour les adapter aux situations de diffusion.

La Figure 10 présente l'évolution de la masse sonore, où la texture (la couleur), soit la troisième dimension, est représentée par l'énergie du signal audio ; et la courbe au tour de 2000 Hz représente l'évolution de la variance spectrale (soit l'étendue spectrale).

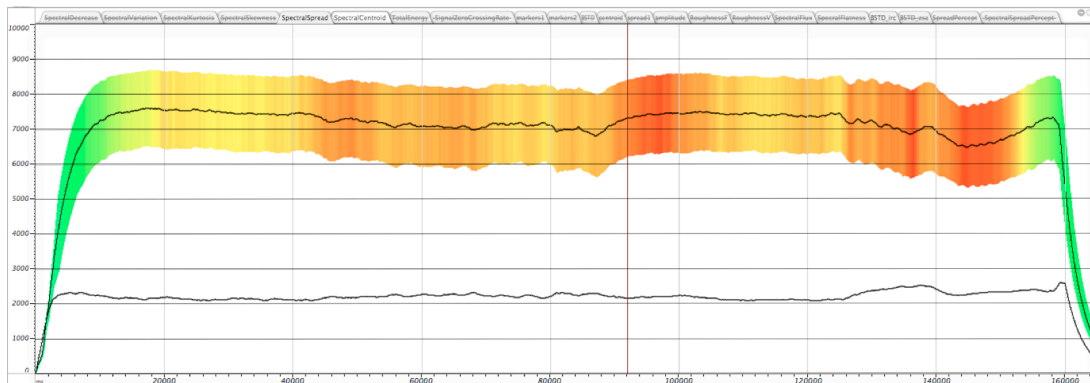


Figure 10. L'évolution de la « masse sonore » et de la variance spectrale de *Concret PH* (1958).

Première constatation : l'aspect « monobloc » de la représentation. Le centroïde varie peu¹⁶, avec une évolution au tour de 7000 Hz, et une amplitude de +/- 250 Hz. La variance spectrale est aussi un paramètre qui varie assez peu. Il présente un mouvement légèrement ascendant de 2200 Hz à 2500 Hz sur la durée de la pièce. De ces deux paramètres, le centroïde montre de petites inflexions, cependant insuffisants pour inférer une segmentation.

Comme nous le rappellent Harley et Valsamakis, l'unique matériau sonore utilisé dans *Concret PH* est issu de l'enregistrement de craquements et sifflements de charbon de bois en

¹⁶ Il va de soi que « varie peu » est une notion qualitative qui devrait être étayée par des mesures telle que la variance, etc. Cela est le sujet d'une prochaine recherche qui est en cours.

combustion. Les techniques de manipulation étant limitées à des manipulations simples de la bande comme l'épissage, la superposition et la transposition, cela à amené Xenakis à la composition d'une œuvre fondée sur la juxtaposition de plusieurs couches de matériau sonore afin d'obtenir différents degrés de densité sonore [Valsamakis 2009, 23]. En plus, « l'étude de la densité et de l'application de fonctions stochastiques pour le calcul de points d'articulation pour chaque couche du son a été une des principales préoccupations. » [Harley 2002, 37], de Xenakis dans cette œuvre. Avant de continuer, et à la lumière de ces faits, nous pourrions inférer ce que Agostino Di Scipio [Di Scipio 1997, 168] nome de propriétés autosimilaires, dans *Concret PH*, découle du processus de manipulation/composition, soit du copier/coller de plusieurs « morceaux de bande ».

Après ces constatations, nous pouvons nous rendre compte qu'il est nécessaire de changer de perspective. La représentation de la « texture », dans le sens donné pour la représentation de la masse sonore, par l'amplitude ou l'énergie du signal, n'est pas dans ce cas utile. D'après les témoignages de Harley et Valsamakis, un des paramètres explicites dans la composition de *Concret PH* est le mouvement de la masse, ou de la texture sonore, il serait peut-être intéressant d'utiliser pour la texture un descripteur permettant de nous donner une information au sujet de la densité sonore. Dans ce cas précis, nous avons utilisé l'évolution de la « variation spectrale perceptive » (Perceptual Spectral Variation).

La Figure 11 présente l'évolution de la masse sonore, ayant comme troisième dimension (la couleur) la variation spectrale perceptive. Dans cette figure nous avons doublé l'information concernant la variation spectrale perceptive. Elle apparaît comme la troisième dimension dans la représentation d'évolution de la masse sonore, et comme courbe. Cela pour souligner son importance. Ce qu'on peut remarquer est le fait que ce descripteur apporte un point de vue nouveau sur l'articulation de *Concret PH*. Si l'étendue spectrale reste moyennement statique, la texture, quant à elle présente une évolution du point de vue de sa densité.

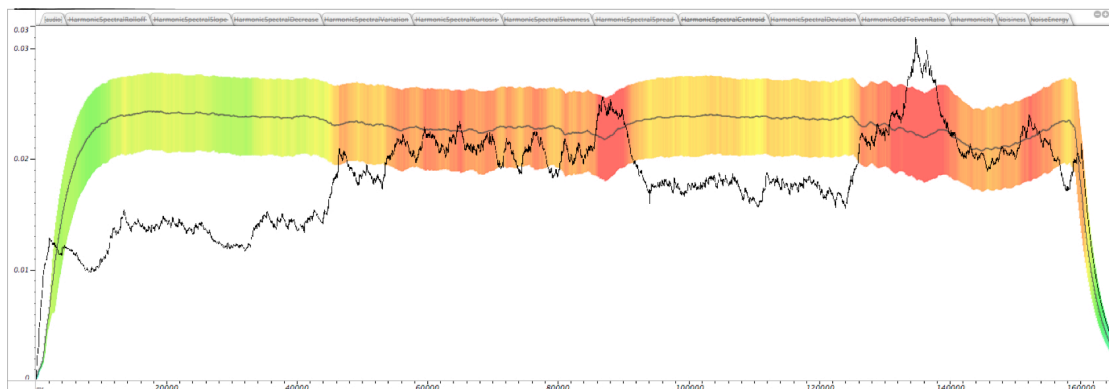


Figure 11. L'évolution de la « masse sonore » et de la variation spectrale perceptive de *Concret PH* (1958).

En nous fondant sur l'évolution de la variation spectrale perceptive, nous proposons une segmentation de *Concret PH*, (voir Tableau 3 et Figure 12) et une forme en quatre parties, qui nous retournent une forme en quatre parties s'appuyant sur l'évolution et le contraste de densités de mouvement à l'intérieur de la masse sonore.

Section	Segment	Position (ms)	Commentaires	Fonctions musicales
I	1	0 → 28730	Stabilisation de la texture, image frontale	Introduction, présentation du matériau, évolution graduelle de la densité sonore perçue
	2	28730 → 43157	Entrée du premier son « grave », Enrichissement du matériau	
II	3	43157 → 53145	Rythmes et début des canons entre les canaux gauche et droit. Mouvement interne	Changement de texture de la masse sonore. Développement, avec diminution de la tension pour préparer à la prochaine section
	4	53145 → 91609	Stabilisation de la texture – fin des canons Répétitions de microformules rythmiques	
III	5	91609 → 123420	Texture légèrement plus aiguë – diminution du mouvement interne Processus de filtrage, évolution de la texture	Évolution de la texture, Contraste de la densité interne du matériau. Point culminant.
	6	123420 → 142051	Début de cinq séquences répétées. Augmentation du mouvement interne.	
IV	7	142051 → 156217	Texture évolue légèrement vers le grave,	Coda
	8	156217 → 164370	Stabilisation — Geste de fin	

Tableau 3. Proposition de segmentation pour *Concret PH*.

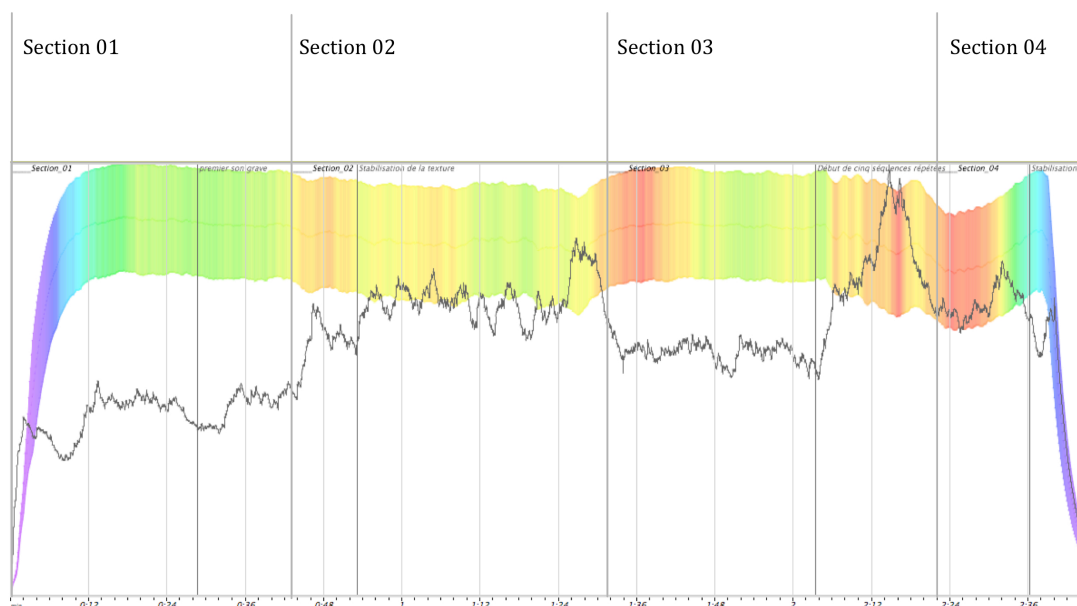


Figure 12. Segmentation à partir de l'évolution de la variation spectrale perceptive de *Concret PH* (1958).

Si la variation spectrale perceptive a pu nous aider en nous donnant des pistes pour la segmentation, d'autres descripteurs peuvent aussi nous être utiles en nous proposant des évolutions d'autres paramètres. Les figures ci-dessous nous présentent l'évolution de la rugosité (Voir [Prestnitzer & al. 1996]) et de la « spectral flatness » (la « planéité spectrale »).

La rugosité psychoacoustique [Prestnitzer & al. 1996] nous permet d'évaluer, ou d'avoir une idée de l'évolution de la tension liée aux composants spectraux d'un son (Figure 13) nous confirmant une conduite (écriture) du matériau sonore qui confirme l'articulation proposée ci-dessus. Dans la première section, nous observons une montée de la rugosité, avec une diminution et reprise à la fin de cette section (la I). Dans la deuxième section, un mouvement décroissant de la rugosité nous prépare au pic de la troisième section. La troisième section,

après le pic, présente une décroissance de tension pour remonter et décroître à la quatrième section, la coda dans la troisième section.

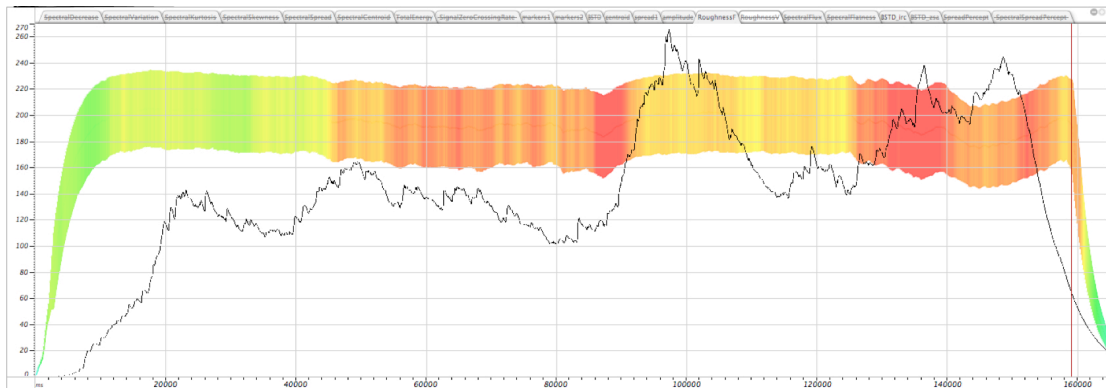


Figure 13. Evolution de la rugosité psychoacoustique de *Concret PH* (1958).

Quant à la « spectral flatness » ou « planéité spectrale » (Figure 14), elle peut être vue comme une mesure de la quantité de bruit d'un signal ou, de manière inversée, de la « tonalité ». Si la « planéité spectrale » est proche de zéro nous avons à faire à un spectre avec des pics, plus « tonale » du point de vue acoustique. Si la valeur s'approche de « un », nous avons à faire à un spectre plutôt plat, soit plus proche d'un bruit. De ce point de vue *Concret PH* peut-être aussi pensé comme une forme fondée sur l'augmentation de l'entropie du matériau sonore.

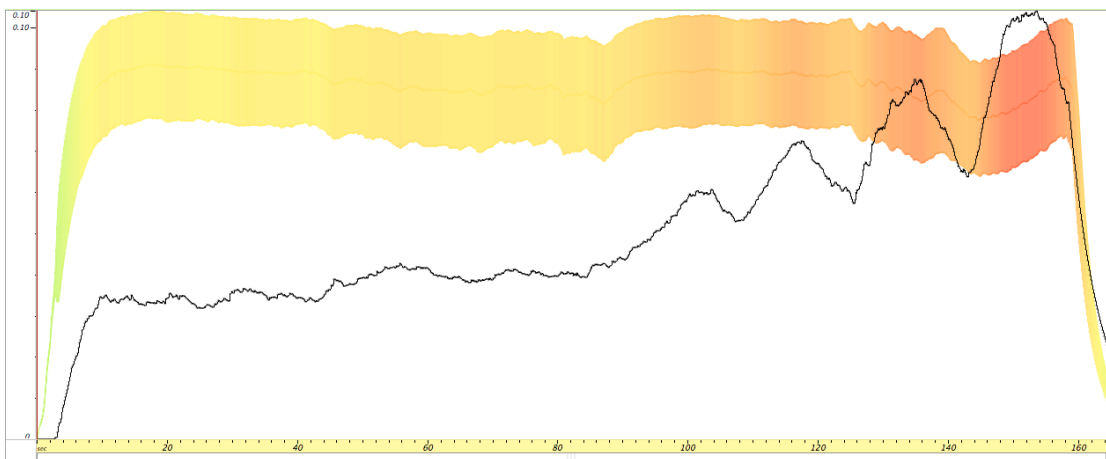


Figure 14. Evolution de la « planéité spectrale » de *Concret PH* (1958).

Cet exemple d'utilisation de descripteurs audio, dans le cadre de l'analyse de *Concret PH*, nous donne des pistes sur un fait bien connu qu'est la diversité d'interprétations d'un phénomène, en fonction du point de vue adopté pour son étude. Dans le cas de l'analyse et de l'interprétation d'œuvres de Xenakis nous pouvons inférer que la forme musicale est fondée sur une perception brute du fait sonore, manipulant l'énergie contenue à différents niveaux pour créer de tensions internes dans le discours musical.

5. ORIENT-OCCIDENT (1960)

Comme nous le rappelle Harley [Harley 2002, 38], *Orient-Occident* est à l'origine la bande-son pour un film documentaire du même nom d'Enrico Fulchignoni, pour l'UNESCO. Le film présente et compare des œuvres d'art de plusieurs civilisations, en allant de la préhistoire à Alexandre le Grand [Valsamakis 2009, 24]. La version originale de 22 minutes n'étant pas très connue et difficilement disponible, nous nous sommes basés sur la version

« concert » de Orient-Occident [Xenakis 1960] et sur la segmentation faite par Solomos [Solomos 2009a].

Pour cette œuvre, Xenakis a utilisé des matériaux sonores venant de plusieurs sources, de sons produits par des boîtes en carton, plaques de métal, gongs, tam-tams et autres, joués avec un archet de violoncelle. Valsamakis remarque aussi que Xenakis a utilisé un enregistrement de *Pithoprakta* (1955), transposé et avec la vitesse de lecture réduite. En général, le matériau sonore utilisé est moins bruyant que celui utilisé dans ses œuvres précédentes, même si la texture du charbon en combustion apparaît dans la dernière partie de l'œuvre [Harley 2002, 38], et que plusieurs textures granulaires apparaissent à divers moments, par exemple, la texture granulaire à 1 min 12 s, ou à 3 min 29 s en début de la deuxième section.

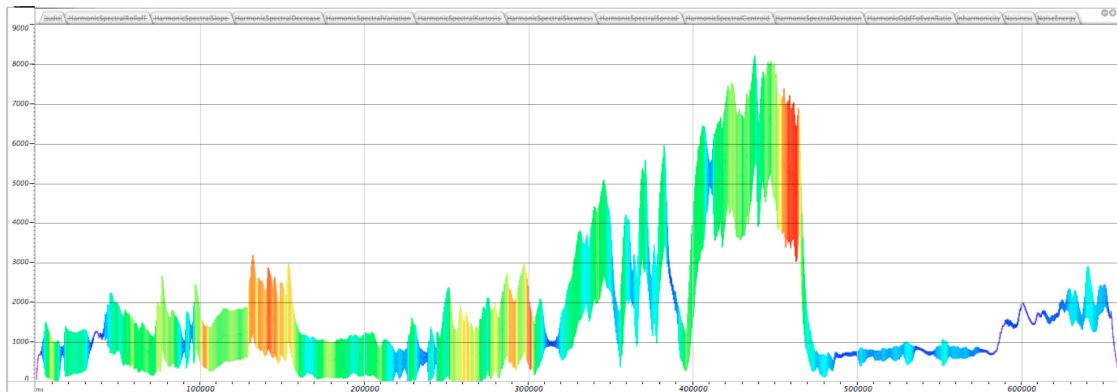


Figure 15. Construction de la représentation d'évolution de la « masse sonore » à partir de la BStD.

La Figure 15 nous présente la représentation de l'évolution de masses sonores. Un premier constat est l'ambitus de variation du centroïde spectral dans cette œuvre qui est de 0 à 8142 Hz, assez étendu, comme pour les œuvres précédentes. Pour Rappel, dans *Diamorphoses* la variation est entre 0 et 7800 Hz, pendant que pour *Concret PH*, le registre spectral reste assez statique autour de 1800 et 2600 Hz, et le centroïde autour de 7000 Hz. Comme nous le verrons par la suite, *Orient-Occident* présente beaucoup de similarités avec *Diamorphoses*.

En revenant à *Orient-Occident*, cette pièce présente une évolution de la masse sonore plus articulée, pourrions-nous dire plus écrite, que les deux œuvres précédentes. Ce que cela veut dire est qu'on perçoit plus d'articulations dans la forme. Dans les trois premières minutes de *Diamorphoses*, on remarque une masse sonore évoluant de manière massive, monobloque, présentant peu ou pas d'articulations. Par contre, dans les trois premières minutes de *Orient-Occident* (Figure 16), nous détectons facilement neuf formes distinctes (de Ia à Ih, voir Tableau 4), avec un registre spectral aussi variable, qu'on pourrait articuler en deux grandes parties, de Ia à Ig, soit de 0 s à 2 min 37 s et Ih de 2 min 37 s à la fin de la section I, soit 3 min 29 sec.

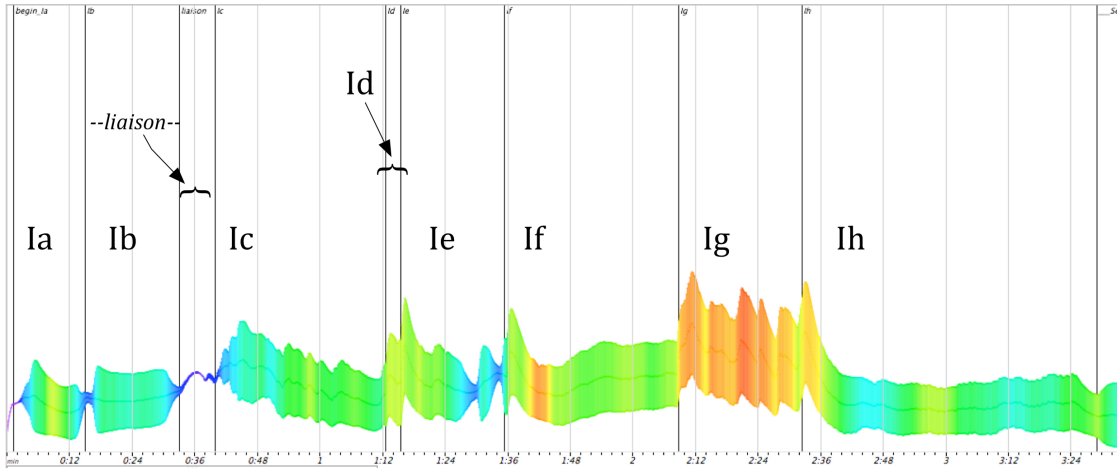


Figure 16. Les neuf formes de masse sonore dans la première section.

Les sous-sections sont :

- Ia, grande masse, geste clairement articulé ;
- Ib, masse, avec une enveloppe comme un souffle, avec un timbre métallique
- Liaison, Un tenu son métallique, avec très faible registre spectral, fait la liaison entre Ib et Ic ;
- Ic, Son métallique de la transition qui s'épaissit, caractère statique, mais avec la couleur qui va vers le grave ;
- Id, courte section, articulée au début par trois attaques consécutives rapides et à la fin par un métallique, comme un fouet ;
- Ie, masse commençant avec un « coup de fouet métallique », un son stable, au tour de 303 Hz, un ensemble d'attaques de percussions et des sons métalliques continus
- If, commence avec un « coup de fouet métallique », présente des sons percussifs attaques presque régulières (apparemment deux boucles qui se superposent avec des coïncidences), sur un fond statique avec une coloration vocale
- Ig, commençant aussi avec le « coup de fouet métallique », développe les éléments de If, avec un accroissement de l'amplitude ;
- Ih, qui commence aussi avec un son ressemblant au « coup de fouet métallique », dilue les éléments présentés antérieurement, en décroissant les rythmes des attaques.

Cette première partie présente une évolution croissante des masses et de la couleur sonore (représentée par le centroïde spectrale) qui a son point culminant à la sous-section Ig, avec une retombée de la tension, de la couleur et du registre spectral en Ih. Cette évolution est soulignée aussi bien par l'évolution du centroïde spectrale, que du registre spectral, qui sont fortement corrélés dans cette première partie.

Sous-section	Position (ms)
<i>Début</i>	0
<i>Ia</i>	1231.98
<i>Ib</i>	14959.8
<i>liaison</i>	33087.5
<i>Ic</i>	39951.4
<i>Id</i>	71735.1
<i>Ie</i>	75476.3
<i>If</i>	95390.6
<i>Ig</i>	128830.
<i>Ih</i>	152414.

Tableau 4. Position des sous-sections de la première section.

Par rapport à *Diamorphoses* le fait d'avoir une évolution des masses sonores plus articulées pourrait s'expliquer de deux manières. La première est de se dire que le compositeur affine son écriture électroacoustique, la complexifie et que la masse s'articule de plus en plus. Peut être, d'autres œuvres, comme *Hibiki Hana Ma* (1969-1970) pourraient affirmer cette hypothèse, tandis que d'autres comme Bohor (1962) ou *La Légende d'Eer* (1977-78) seraient la preuve du contraire. Dans le cas d'*Orient-Occident*, nous pensons que le caractère articulé de l'œuvre est plutôt tributaire de l'aspect visuel avec lequel la musique devait se synchroniser ou dialoguer. La mise en rapport de la représentation de la masse sonore avec les objets montrés dans le film (cf. Solomos 2009a), permet de se rendre compte de la corrélation entre l'articulation de la masse sonore et des événements du film, dont la présentation des différents objets Figure 17).

Dans le Tableau 5 et dans la Figure 17, nous pouvons observer les événements du film relatifs à cette première partie d'*Orient-Occident*. Ces événements étant constitués de l'apparition de nouvelles images, ou de changements de plans visuels.

Evènement	Positions (ms)
"---générique début---	3000.00
Vénus tournante	16000.0
gros plan sur visage-main	39000.0
femme au collier	48000.0
musiciens	71000.0
plan rapproché	77000.0
premier bouddha	106000.
panoramique des pieds à la tête	122000.
bouddha jambes croisées	149000.
panoramique des pieds à la tête	157000.
personnages drapés	174000.
gros plan	192000.
II-partie_ petit lion	209000.

Tableau 5. Evènements de la première partie, selon Solomos [Solomos 2009a].

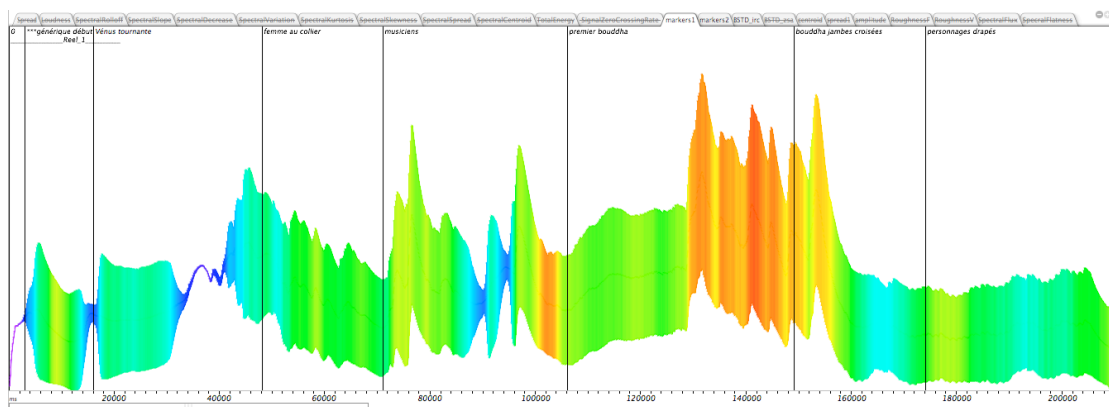


Figure 17. Première section d'*Orient-Occident* et les événements du film.

La mise en relation des Tableau 4 et Tableau 5 nous permet d'observer comment les événements de la partie musicale suivent de manière souple les apparitions de nouveaux objets ou de changements de plans visuels.

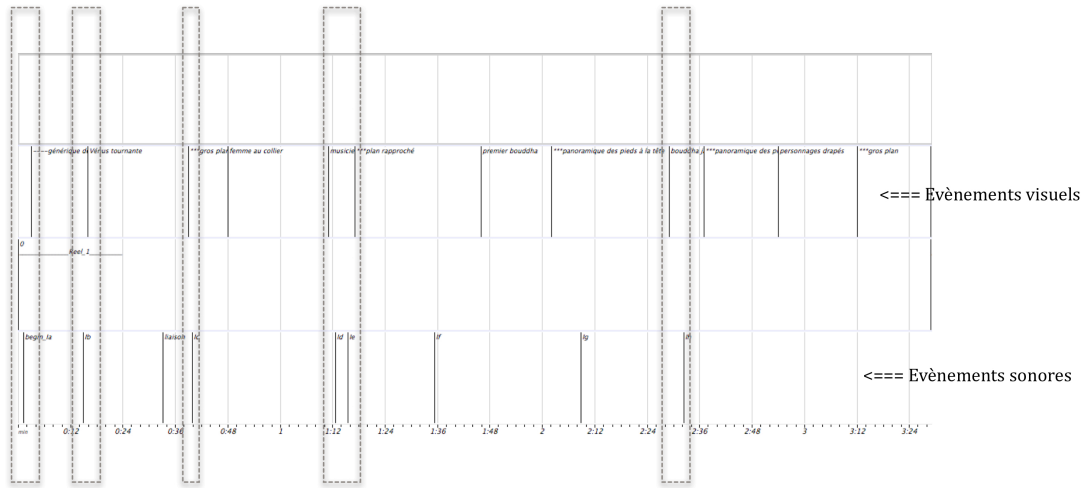


Figure 18. Rapport entre les articulations de l'image et des masses sonores, dans la première partie d'Orient-Occident.

La Figure 18 résume ces rapports de manière graphique. On perçoit aisément l'articulation ou la synchronisation entre le générique du début et la sous-section Ia. La section Ib commençant à 15 s et la présentation de la « Vénus tournante » à 16 sec. Le deuxième point de synchronisation est autour de 39 sec, avec la sous-section Ic et le changement de plan dans l'image, avant la « femme au collier ». L'apparition des musiciens à 1 min 11 s sera suivie de la courte sous-section Id. La sous-section Ie (à 1 min 15 s), précédant de le plan rapproché sur les musiciens à 1 min 17 s. Le prochain point de synchronisation (flou) sera autour de l'apparition de l'image « bouddha jambes croisées » à 2 min 29 s, l'entrée de la sous-section Ih à 2 min 32 s, et le mouvement de caméra sur le « bouddha jambes croisées » (panoramique des pieds à la tête) à 2 min 37 s. Pour en finir à 3 min 29 s, le début de la deuxième section.

Même si on constate la corrélation entre la partie musicale et la partie visuelle, on peut aussi constater que pour cette version de concert, la partie musicale présentera sa propre logique, indépendante de la séquence visuelle.

Pour finir avec cette première section, nous pouvons aussi observer une évolution croissante de la variation spectrale perceptive (Figure 19). Cette courbe corrobore le fait que cette première section présente une évolution croissante de la matière musicale, du mouvement interne des textures et de la masse sonore. Nous pourrions inférer que cette première partie se compose d'un mouvement croissant d'énergie, clos par une section conclusive ayant aussi une fonction de transition à la deuxième section.

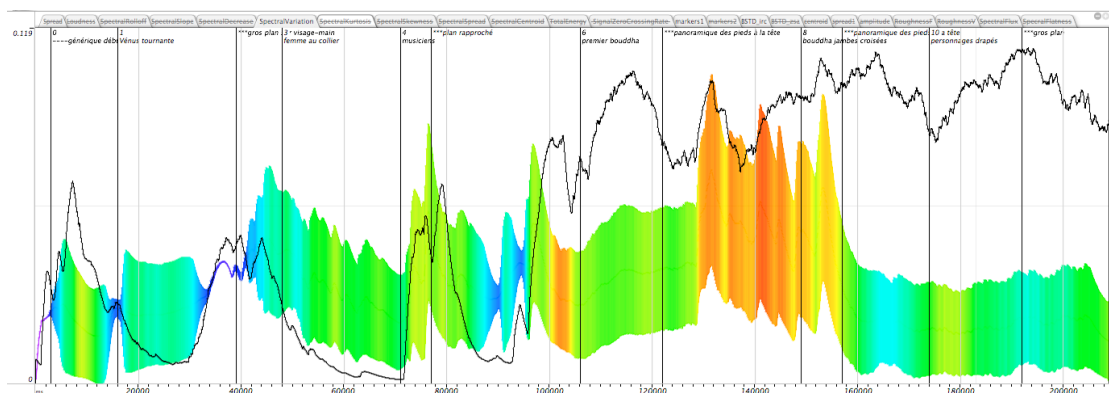


Figure 19. La première section d'Orient-Occident avec la courbe de Variation spectrale.

Un autre constat est le fait que la représentation des masses sonores, avec la BSTD, s'aligne avec la segmentation des « trois bobines » proposée par Solomos [Solomos 2009a], de

manière que la première partie, correspondant à la première bobine, aille de 0 s à 3 min 29 s, la deuxième partie (la deuxième bobine) aille de 3 min 29 s à 7 min 46 s, et la troisième partie de 7 min 46 s jusqu'à la fin (10 min 58 s).

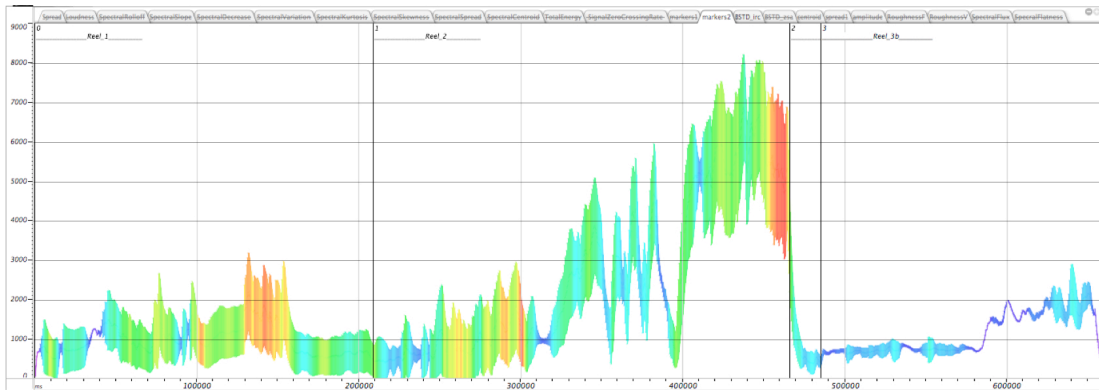


Figure 20. Alignement de la représentation avec la segmentation proposée par Solomos [Solomos 2009a].

La deuxième section présente aussi une articulation claire, que nous présentons à la Figure 21 et par le Tableau 6.

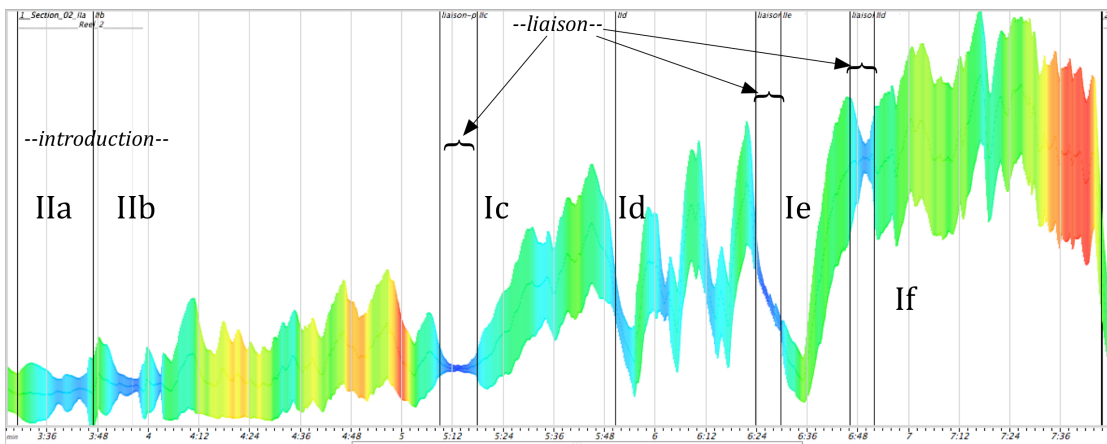


Figure 21. Articulation de la deuxième section d'*Orient-Occident*.

Sous-section	Position (ms)
<i>Ila</i>	209000
<i>Ib</i>	226997
<i>liaison</i>	309131
<i>Ic</i>	317849
<i>Id</i>	350657
<i>liaison</i>	383759
<i>Ile</i>	389888
<i>liaison</i>	406237
<i>If</i>	411834
<i>Section 03</i>	466000

Tableau 6. Position des sous-sections de la deuxième section.

Dans cette deuxième section, nous observons une nette augmentation de la couleur et du mouvement interne de la masse spectrale. La sous-section IIa, prolonge la sous-section Ih, et sert d'introduction à de nouveaux matériaux présentés à IIb. IIb propose une masse sonore évoluant d'un timbre grave, foncé à une couleur plus métallique. À 5 min 9 s, une courte

section, sur un son métallique tenu avec un centroïde spectral assez stable, soit autour de 946 – 957 Hz, et un registre spectral, très mince au tour de 92 Hz, fonctionne comme une liaison, un pont, une transition à la sous-section IIc. IIc présente une masse sonore avec une évolution nette et marquée de la couleur, avec un centroïde évoluant de 950 à 3700 Hz et un registre spectral passant de 92 Hz à 2300 Hz. La sous-section II d se caractérise par une oscillation de timbres, ou couleurs sonores réalisés par le contraste et opposition marquée de matériaux, plus que par une interpolation ou transition lente. Cette sous-section se caractérise aussi par un registre spectral plus variable évoluant entre 1600 Hz et 2700 Hz. Une courte sous-section de liaison avec IIe qui comme IIc propose une masse sonore évoluant, de manière marquée, de matériaux graves, foncés à des matériaux plus bruiteux, plus brillants, avec un centroïde évoluant de 700 à 5000 Hz. Une nouvelle sous-section de liaison, présentant une texture granuleuse pour nous amener à II f, sous-section massive présentant un centroïde spectral haut, soit entre 4700 et 6880 Hz, qu'est le point culminant de cette œuvre.

La troisième section s'articule clairement en trois sous-sections. De 7 min 46 s à 8 min 5 s, une petite section de transition, IIIa et IIIb. Chacune de ces sous-sections présente deux sous-sections, clairement visibles.

Cependant, les deux sous-sections IIIa et IIIb diffèrent et contrastent par la nature des matériaux utilisés et de la couleur. IIIa présente des matériaux nettement harmoniques et tonals (dans le sens de hauteur clairement perçue), avec une sensation générale d'immobilité et de distance magnifiée par la réverbération. Dans IIIa le centroïde spectral est assez stable au tour de 700 Hz.

IIIb présente un matériau plus bruité, granuleux, aussi réverbéré, mais avec plus de mobilité et une couleur plus « claire » que IIIa, avec un centroïde spectral plus mobile (que dans IIIa) autour de 1700 Hz.

Par rapport à la pièce, de manière globale, cette partie (la section 03) se présente de manière plus stable, presque immobile, nous proposant une fonction musicale de coda en deux parties.

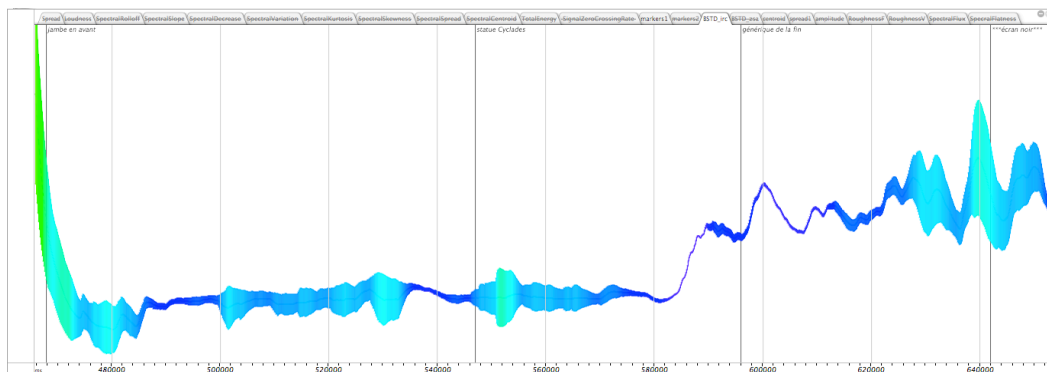


Figure 22. Pédale de fin en deux parties.

Pour finir, voici une proposition d'articulation d'*Orient-Occident* présentée dans le Tableau 7.

Section 1		Section 2		Section 3	
De 0'00" à 3'29"		De 3'29" à 7'46"		De 7'46" à la fin 10'56"	
<i>Ia Ib liaison Ic Id Ie If Ig</i>	<i>Ih</i>	<i>Ila Ib liaison Ic Id liaison Ile liaison If</i>		<i>Introduction IIIa IIIb</i>	
Mouvement général ascendant		Déploiement de l'énergie et des couleurs		Stabilité en deux paliers - Coda	

Tableau 7. Résumé formel pour *Orient-Occident*.

6. UNE PREMIERE COMPARAISON - *DIAMORPHOSES/ORIENT-OCCIDENT*

Une des premières utilisations que nous pouvons faire de la représentation de la « masse sonore » est de pouvoir, très rapidement, avoir un aperçu de l'articulation globale de l'œuvre en étude. Par exemple, si nous mettons côté à côté *Diamorphoses*, (durée 6 min 56 s et évoluant entre 0 et 8300 Hz) et *Orient-Occident* (durée de 10 min 58 s et évoluant entre 0 et 8100 Hz), la représentation graphique nous autorise à nous affranchir des rapports temporels entre les deux pièces pour les afficher à la même échelle (Figure 23), une sorte de « normalisation temporelle ». Cette stratégie permet de se rendre compte des ressemblances (et différences) de l'évolution globale de la masse sonore des deux représentations.

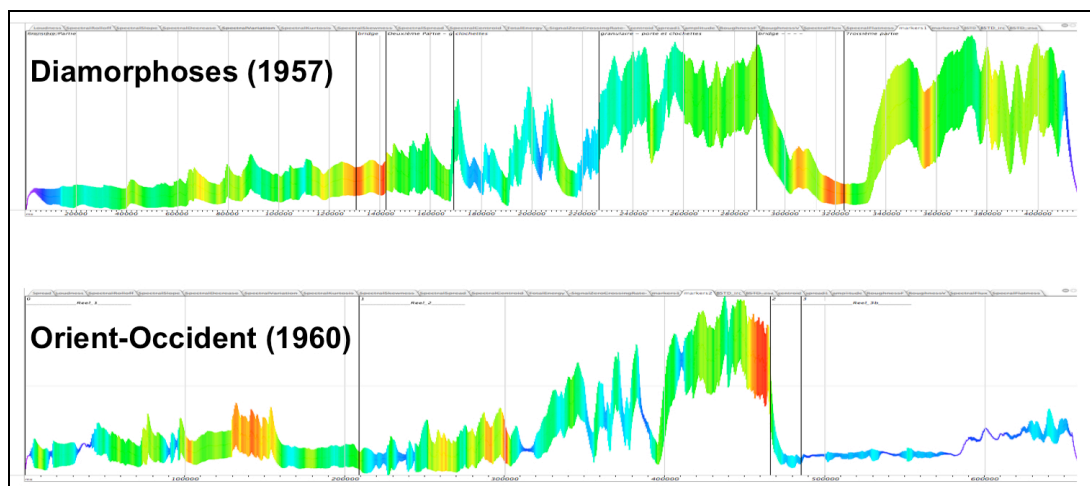


Figure 23. Superposition des représentations de « masse sonore » et normalisation temporelle.

En nous basant sur les catégories proposées dans le Tableau 1, nous pouvons remarquer la grande similarité des deux œuvres (Figure 24). Pour cette comparaison, nous avons considéré, pour les sections centrales, un mouvement ascendant global des diverses morphologies, obtenant une articulation globale des catégories : {1 3} {1 2 1 3} {2 2}, ce qui nous retourne une structure tripartite, A-B-Coda, et deux « ponts » (représentés par des morphologies de la catégorie 3) entre les sections. Dans les deux cas, un point culminant (des masses sonores) se présente à la fin de la deuxième section, les sections finales présentant des textures, des masses sonores « statistiquement statiques », avec une fonction vraisemblablement de « pédale », de coda à la pièce (Figure 25).

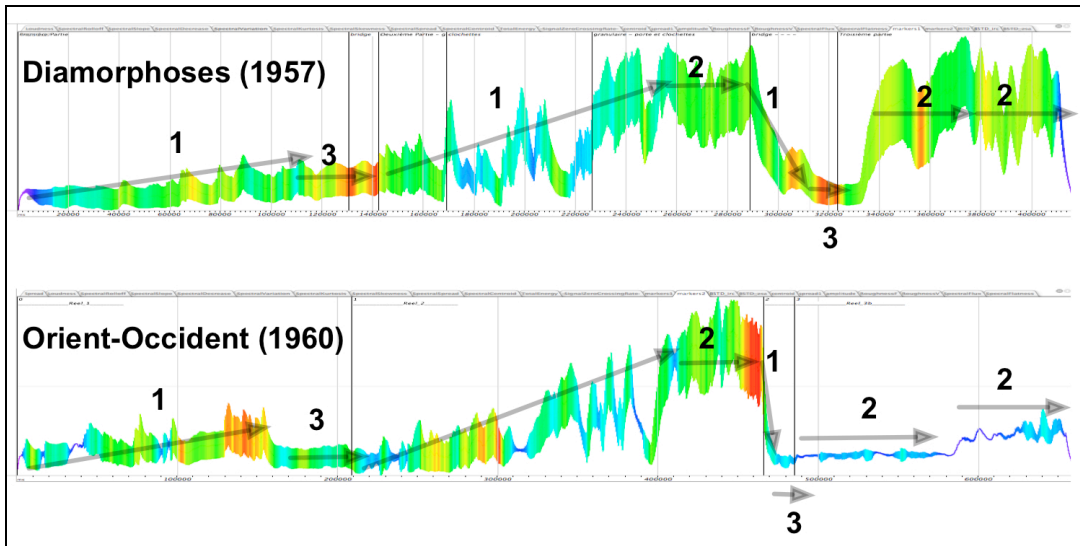


Figure 24. Superposition des représentations et catégories.

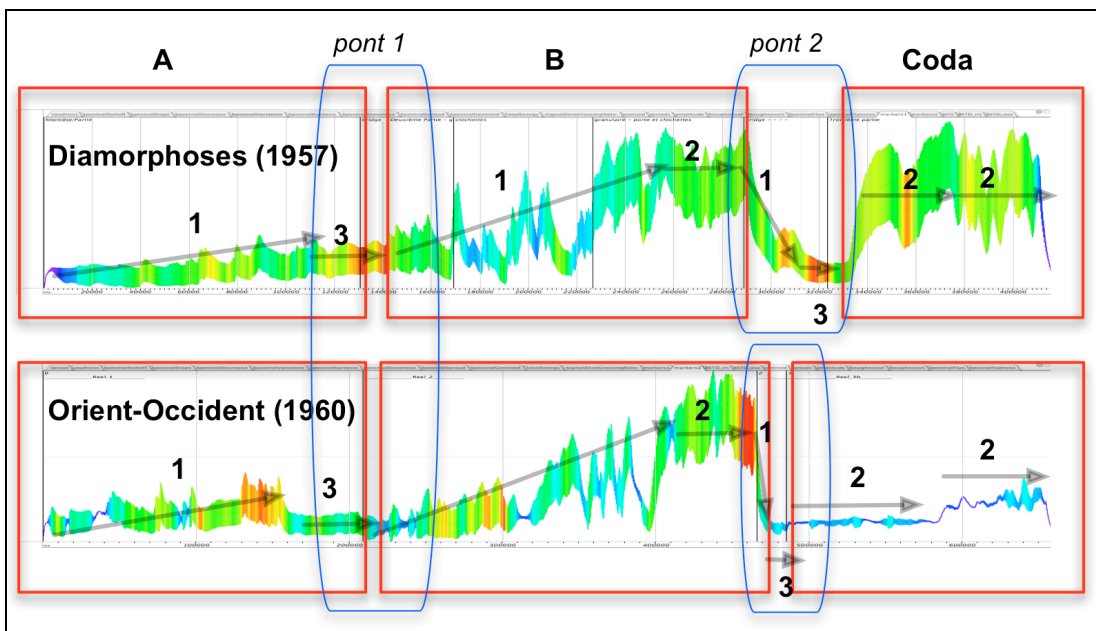


Figure 25. Comparaison des articulations formelles.

Les deux œuvres présentent, globalement, la même forme tripartite, avec la même conduite des masses sonores. Section 1 évolution croissante de la masse sonore, avec une dernière sous-section plus statique préparant l'entrée de la deuxième section. Section 2, mise en opposition de sous-sections avec une lente évolution de la masse sonore avec des sous-sections présentant de larges mouvements dans la couleur. Section 3 en deux parties, précédées d'une courte partie transitoire (un pont) avec tendance à l'immobilité.

La grande différence réside dans le fait que dans *Orient-Occident*, les sections se présentent plus articulées et plus détaillées.

7. BOHOR (1962)

L'analyse de Bohor, du point de vue de la méthodologie que nous proposons, présente trois problèmes principaux : l'analyse d'une œuvre multicanal à partir de l'aplatissement des canaux individuels, la question du bruit de fond de certains enregistrements (notamment sur les fichiers originaux) et finalement la question de la mise en relation des diverses informations retournées par les descriptions audio. Dans ce contexte, nous sommes en train de

travailler sur deux enregistrements pour notre analyse. L'enregistrement commercial sorti en disque compact de 1997 [Xenakis 1997d] et les quatre originaux numérisés correspondant à la deuxième version de l'œuvre [Couprie 2007]¹⁷. *Bohor* est décrite comme étant une œuvre huit pistes. Comme le remarque Pierre Couprie [Couprie 2007], ces huit pistes peuvent être réduites à quatre pistes, vu que les fichiers stéréo originaux contenaient le même matériau sur les pistes de droite et de gauche. Ce sont ces quatre fichiers sur lesquels nous sommes en train de travailler.

Dans le cadre de cette communication, nous présenterons seulement une analyse sommaire de l'aplatissement des deux canaux stéréo de la version commerciale de 1997.

Comme nous l'avons remarqué, notre analyse de *Bohor* aurait pu souffrir d'un problème méthodologique, par l'aplatissement des huit canaux. Selon les analyses de Pierre Couprie [Couprie 2007], de Valsamakis [Valsamakis 2009] et de Gibson [Gibson 2008], nous pouvons nous rendre compte que les différentes pistes possèdent des matériaux et des évolutions de densités différentes. La disposition spatiale des différentes pistes étant aussi sujettes à discussion et à l'étude.

Nous sommes conscients qu'il existe une différence fondamentale du point de vue de la perception d'écouter l'ensemble des pistes émanant du même point spatialement, que de les écouter placées, physiquement dans l'espace à des endroits différents. La séparation spatiale de sources sonores aide la ségrégation auditive nous permettant de mieux identifier et séparer les différents événements spatiaux [Bregman 1990, 73-83].

Pour cette version « aplatie » de *Bohor*, nous avons changé les paramètres d'analyse. Nous utilisons une taille de fenêtre de 32 768 échantillons et un pas d'avancement de 16 384 échantillons, correspondant à 743,04 ms et 371,52 ms. Nous justifions ce choix par le fait d'analyser une pièce avec une évolution lente, et que notre approche est ici globale. Une analyse plus fine se justifierait si nous présentions les fichiers séparés.

Avec une amplitude de variation du registre spectral de la masse sonore, entre 0 et 3714 Hz, la représentation d'évolution de la masse sonore présente une morphologie mono bloque, qui donne très peu d'informations, à première vue, pour la segmentation (Figure 26).

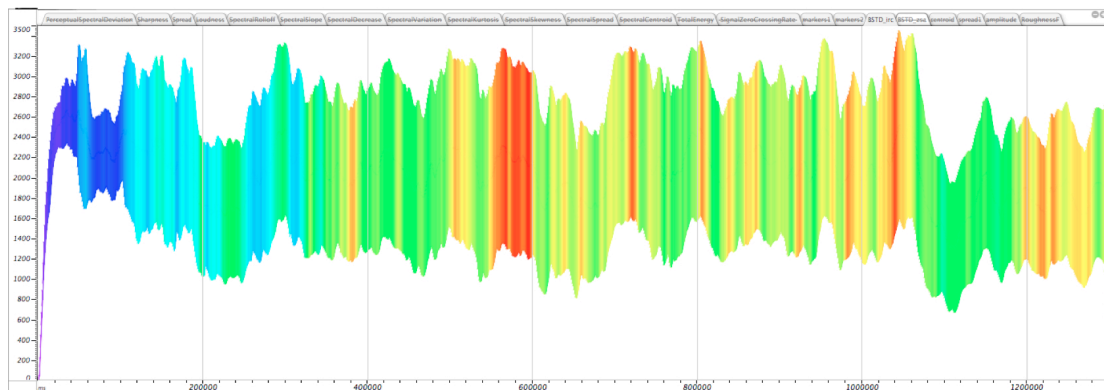


Figure 26. Evolution de la « masse sonore » dans *Bohor* (1962).

Selon [Couprie 2007], la forme globale de *Bohor* est tripartite : A1-A2-B. Le Tableau 8 présente cette articulation.

¹⁷ Nous remercions Jacques-Diego Losa et Daniel Teruggi pour nous avoir donné l'accès à la version originale de *Bohor* en huit pistes réalisée au GRM par Xenakis. Ces quatre fichiers mono étaient des fichiers en format «sounddesigner 2», à un taux d'échantillonnage de 96 KHz et avec une résolution de 24 bits.

Sections	I (A1)	II (A2)	II→III	III (B)
Début-fin	0 – 6'00''	6'00'' – (18'13-18'54'')	(18'13-18'54'')	(18'13-18'54'')-21'31''
Durée	300''	12'13''-12'54''		
Obs.	Mise en place du matériau avec variations importantes d'intensité et de densité	Stabilité et développement du matériau	Partie mobile	Bruit Blanc

Tableau 8. Résumé analytique selon Couprie [Couprie 2007].

La partie que va de 18'13 à 18'54'' est, selon Pierre Couprie [Couprie 2007], « une partie mobile dans sa perception puisqu'elle correspond à l'arrivée du bruit blanc qui ne se trouve pas au même endroit selon les groupes de pistes (1 — 2 : 18'13'', 5 — 6 : 18'54'', 7 — 8 : 18'02''), il s'opère donc un tuilage avec un effet spatial. » Ce qui signifie que l'articulation formelle dépendra fortement de la position de l'auditeur et aussi de la diffusion de la pièce. Dans la représentation de la masse sonore des Figure 26 et Figure 27, cette partie correspond aussi à une baisse de l'énergie (la couleur verte), une baisse de l'étendue spectrale et aussi du centre de gravité de la forme (Figure 27). Pour comprendre le pourquoi de cette zone mobile, il est impératif d'écouter au minimum la version quadriphonique, qui permet de se rendre compte de l'arrivée décalée du bruit blanc sur les différents canaux.

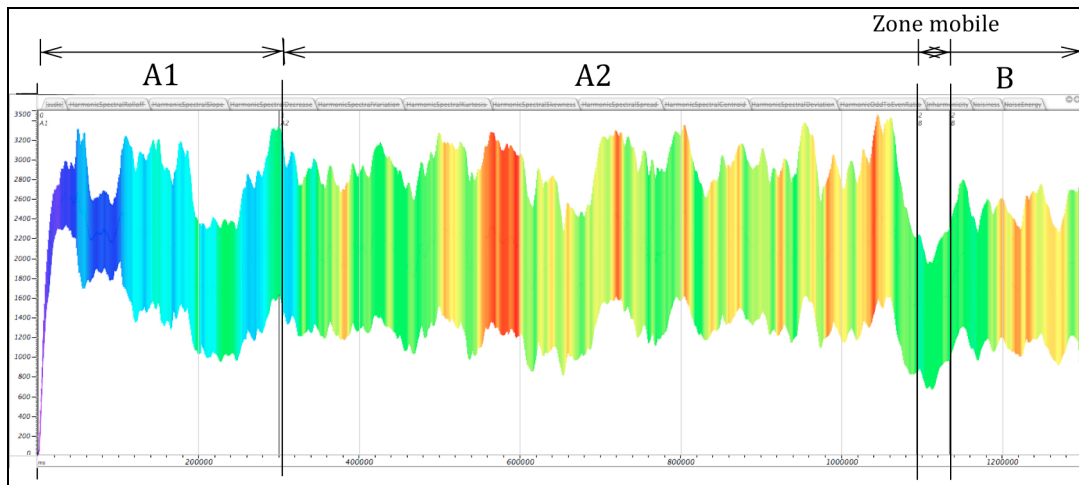


Figure 27. Résumé analytique selon Couprie [Couprie 2007], par rapport à l'évolution de la masse sonore.

En nous fondant sur cette analyse, la représentation de la masse sonore pourra nous aider à compléter l'information. Par exemple, nous remarquons que entre 190 s et 245 s la figure nous présente une rapide décroissance du centroïde spectral et aussi de l'étendue spectrale. En suite, nous remarquons que cette partie est suivie par une ascension de la masse sonore, juste avant le début de la section A2.

À partir de la position 190 s jusqu'à 245 s, nous remarquons aussi, d'un point de vue auditif, un changement de couleur, la perte de brillance (perte d'aigus), une mise en retrait des sons métalliques, soit un changement d'espace et une diminution de la densité d'évènements. Une mise en retrait, un arrêt du mouvement.

Cette partie, de 190 s à 245 sec, et celle qui suit, de 245 s à 300 sec, peuvent être considérées comme une partie de transition (Figure 28), avec l'arrêt du mouvement et de la densité, avec une interpolation de textures entre les sections A1 et A2 (Figure 28), allant

d'une masse sonore sombre (de 190 s à 245 sec) à une masse sonore plus brillante (entrée de A1).

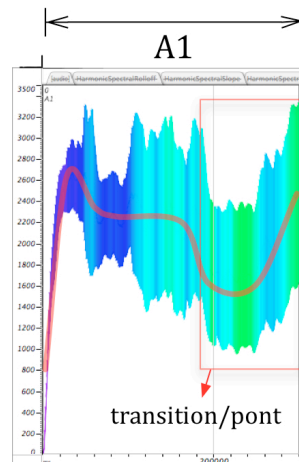


Figure 28. Le pont entre les sections A1 et A2.

Comme nous le signalons au début de cette partie, *Bohor* nous propose un certain nombre de problèmes méthodologiques. Comme nous aimerions plutôt décrire le comportement global, il peut être utile de « lisser » les valeurs d'analyse de manière à obtenir des courbes avec moins d'aspérités, ce qui nous permet d'avoir une meilleure description du comportement globale de la courbe. À partir de l'analyse que nous venons de faire, nous avons lissé les valeurs des descripteurs (avec un filtre passe-bas) pour éliminer les variations locales.

Premièrement, une analyse de l'évolution de l'énergie présentée à la Figure 29, où nous avons aussi signalé la sous-section de transition de A1 à A2. Une constatation est que l'énergie dans *Bohor* évolue avec une forme proche d'un arc comme nous le montre la Figure 30.

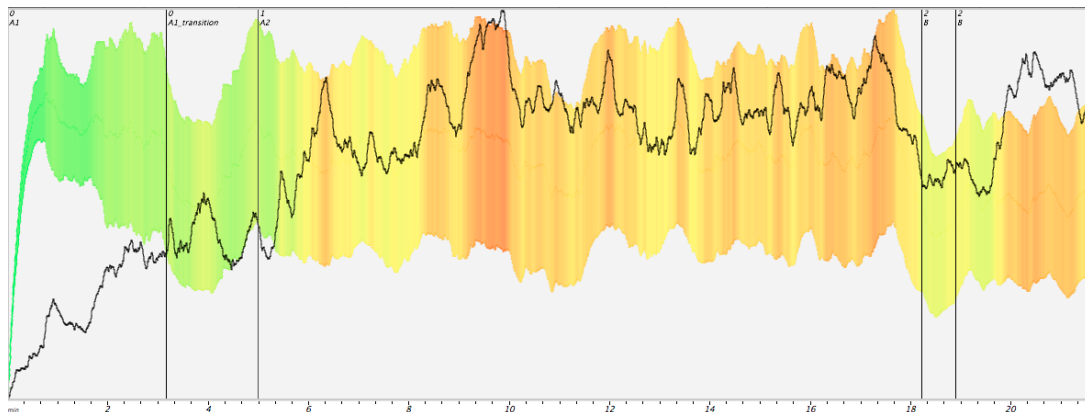


Figure 29. Evolution de l'énergie totale dans *Bohor*.

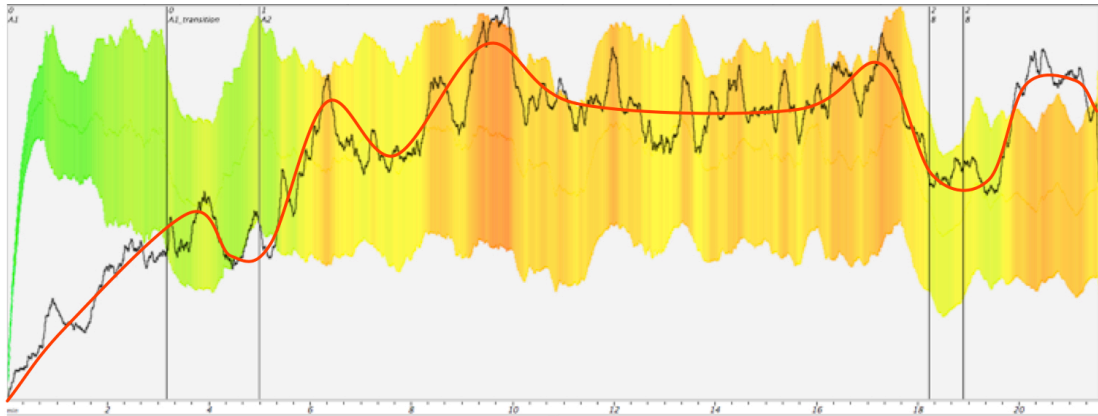


Figure 30. Lissage de la courbe d'évolution de l'énergie.

Ce paramètre (l'énergie) nous montre une constante dans le travail de Xenakis, qu'est l'articulation de la forme et de l'évolution de la matière sonore pilotées par l'intensité sonore, par l'énergie.

La partie centrale (A2) présente des discontinuités importantes du point de vue de certains descripteurs, comme l'énergie, le centroïde spectral, la variation spectrale, etc. Mais, de notre point de vue, il est important de pouvoir séparer les différentes voix (canaux) pour une analyse plus approfondie.

Cependant, il est aisé de remarquer qu'à partir de 17 min 40 s, le début d'un changement important de la couleur se produit. Ce fait est bien signalé par une baisse du centroïde spectral, que de 2500 Hz (à 17 min 38 s) passe à, à peu près, 1500 Hz (à 18 min 30 s). Une variation de 1000 Hz en 52 sec. a nouveau, nous pouvons constater une courte sous-section de transition commençant autour de 17 min 40 s, jusqu'au milieu de la partie mobile (18 min 30 s).

Si on peut avoir une première impression d'œuvre monobloc, statique, nous percevons que *Bohor*, en entier, présente des évolutions structurées de plusieurs paramètres sonores. D'un point de vue global, l'évolution de la « noisiness », le caractère bruité du signal sonore (voir Figure 31), signale une évolution lente, continue et croissante du caractère bruité de la masse sonore.

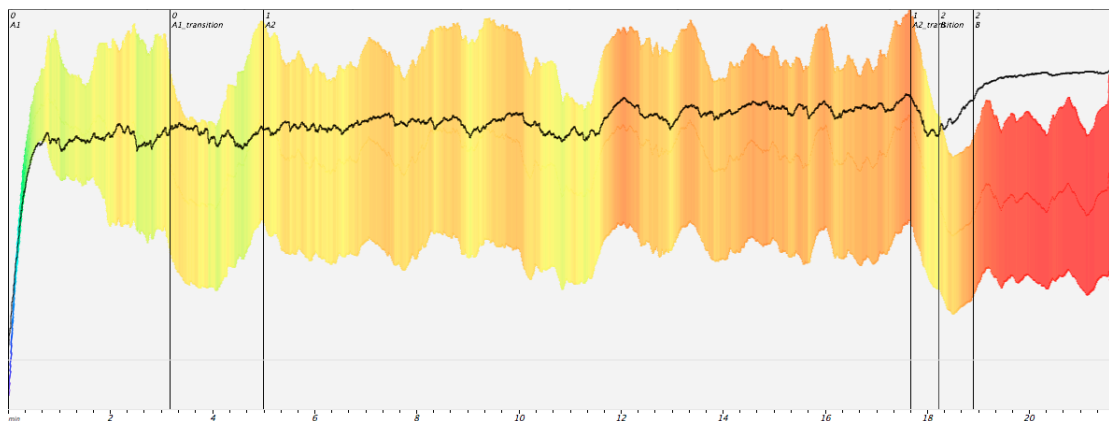


Figure 31. Evolution du caractère bruité (noisiness), la couleur est la « noisiness ».

Avant de finir cette partie dédiée à *Bohor*, nous allons faire une halte à la troisième section. Ce qui est certain est l'apparition, encore une fois d'une morphologie stable sur la durée dans la dernière section. À partir de 18 min 13 s surgit un matériau bruité, fondé sur des transpositions de bruit blanc, et qui restera constant jusqu'à la fin de la pièce. Plusieurs autres descripteurs, comme la variation spectrale, la « planéité spectrale » (« spectral flatness ») ou le « spectral crest » (la crête spectrale), confirment le fait qu'un signal bruité se stabilise à partir de 18'54'', indiquant une articulation claire, avec ce qui précède. Comme on peut le

constater, cette stabilisation concorde avec la fin de la partie mobile, transition entre les sections A2 et B. La présence du bruit blanc est clairement repérée.

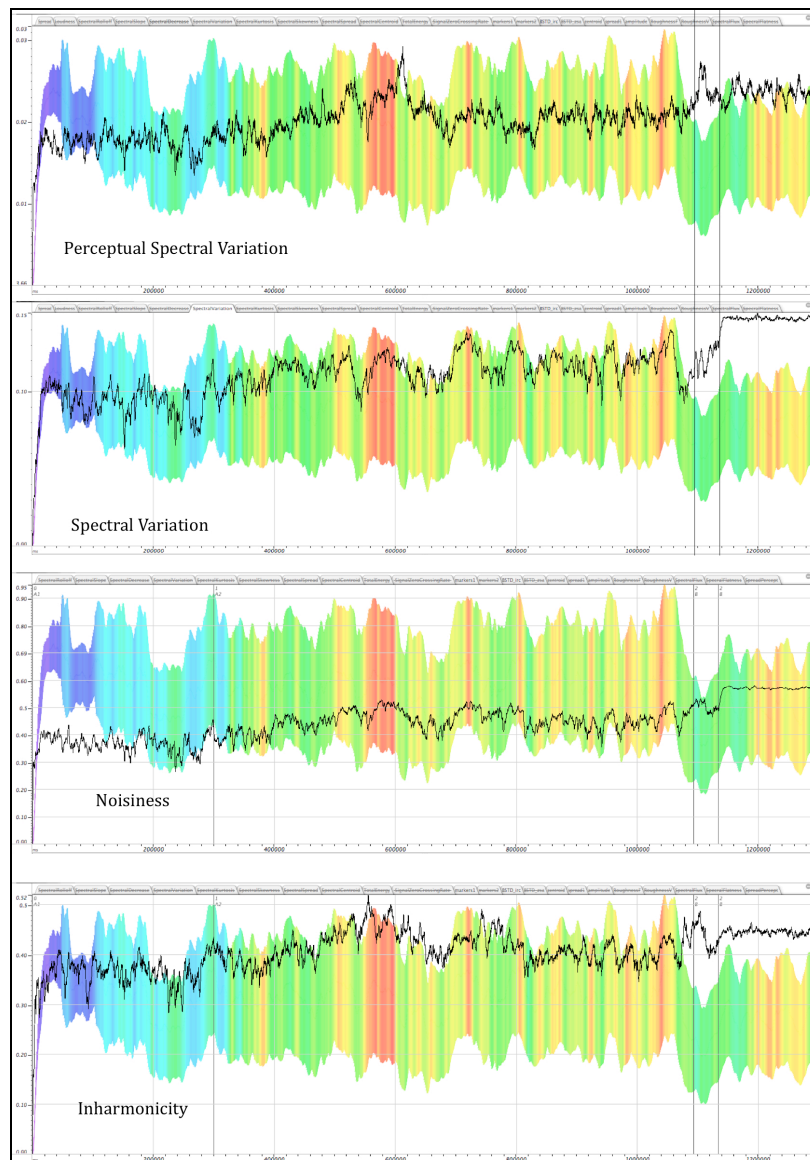


Figure 32. Information auxiliaire apportée par d'autres descripteurs, pour la dernière partie de *Bohor*.

En plus, si nous revenons aux descriptions que nous avons faites, notamment celle de l'énergie, il est possible de se rendre compte que cette dernière section pourrait aussi être segmentée en deux parties (Figure 33). À partir de 19 min 49 s, l'énergie du bruit blanc passe de -23 dB (19 min 34 s), à -20 dB (19 min 58 s), avec des pics à -19 dB. Cet accroissement de l'énergie magnifie les fréquences plus aiguës, donnant une sensation d'éclaircissement, de brillance de la masse sonore.

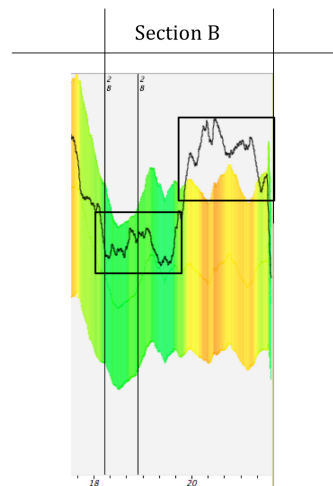


Figure 33. Articulazione della sezione « B » di Bohor, par l'accroissement d'énergie.

Pour finir, *Bohor* nous fait penser à *Concret PH*. Une grande évolution de la masse sonore et du timbre, avec une forme principalement articulée par la texture de cette masse sonore.

Pour reprendre Pierre Couprie « Même si cette œuvre est souvent analysée comme un son continu, trois parties sont assez clairement identifiables » [Couprie 2007]. Notre analyse le confirme. Cependant, nous pouvons ajouter que même si l'analyse que nous en avons faite est assez sommaire, nous pouvons constater un point. Le fait qu'avant chaque articulation de section, précède une morphologie de masse sonore stable suivie par transition ascendante ou descendante. À la fin de l'œuvre, un dernier bloc, plutôt stable se dessine (le bruit blanc), comme une « pédale statistiquement statique » en deux parties, annonçant la fin, avec une fonction possible de coda.

8. CONCLUSIONS

Un premier enseignement de ces quatre expériences, ou exercices de style, est le fait que cette méthodologie (fondée sur l'utilisation des descripteurs audio et de la représentation BStD comme modélisation bas niveau des masses sonores) propose principalement une approche globale, qui nous semble bien adaptée dans le cas de Xenakis, et à des « musiques de texture ». Cette méthodologie se veut une aide à l'écoute, à l'annotation et à la segmentation. Elle se veut principalement un outil de description du phénomène étudié. Comme description, ou représentation, cet outil sera forcément « incomplet ». Les représentations sont par définition « sélectives ». Comme l'écrit Frassen, « toute représentation est sélective et la sélectivité est essentielle pour ce qui est représenté... » [Fraassen 2008, 37]. De ce fait, les informations retournées devront forcément être complétées par d'autres moyens. L'analyse utilisant les descripteurs audio, et toute représentation s'y référant est un point de vue partiel nécessitant un apport d'information.

En plus d'une aide à l'analyse d'œuvres musicales, nous voyons aussi l'utilisation des descripteurs audio comme un outil à la formation d'une oreille. Évidemment, on pourrait se demander si l'outil visuel ne « force » pas une écoute, ne force pas une perception fautive. Si, effectivement, il est possible que le support visuel induise une écoute, comme l'a bien signalé Vincent Tiffon dans sa critique de l'utilisation de l'analyse sonographique [Tiffon 2006]. Il est aussi possible que ce type d'analyse apporte des informations, issues de l'application de l'algorithme, qui ne sont pas directement corrélées avec la perception. Cependant, nous faisons l'hypothèse que ce même support visuel est aussi une aide pour permettre à l'écoute de se focaliser sur une caractéristique sonore, et en conséquence être une aide pour le développement d'un solfège lié au timbre.

Concernant l'utilisation des descripteurs audio, il est aussi important de signaler le problème de la redondance d'information. Plusieurs descripteurs présentent la même information, ou le même apport d'information. Par exemple, la pente spectrale, le « spectral

roll-off» ou le centroïde spectral sont souvent très corrélés. À un niveau global, ils sont souvent interchangeables. Un des premiers défis est d'arriver à choisir le groupe de descripteurs couvrant le plus grand nombre de paramètres musicaux, ou de caractéristiques sonores possibles.

Un autre point, après avoir défini l'ensemble de descripteurs non redondants, est de faire le rapport entre la description scientifique, physique du son et des caractéristiques musicales et/ou auditives. Par exemple, la fréquence fondamentale, dans le cadre de l'analyse de la parole, encode une partie de la prosodie, l'amplitude d'un signal est fortement corrélée avec le phrasé, le centroïde spectral encode une partie de notre perception du timbre, l'écart-type spectral encode une partie de notre perception du « volume sonore », et ainsi de suite. Pourquoi avoir utilisé la « Perceptual Spectral Variation » au lieu de la « Spectral Variation », ou du « Flux Perceptif » ? Ces questions n'ont pas encore de réponses définitives. Un autre point est le fait que les descripteurs de bas niveau (comme ceux que nous avons utilisés) sont insuffisants pour modéliser une perception musicale complexe, raison pour laquelle une méthodologie que nous pourrions proposer est la construction de descripteurs composés, telle que la BStD, que nous avons utilisée pour modéliser le concept de « masse sonore ».

Un autre aspect que nous aimerions pointer est le fait que l'utilisation des descripteurs est très sensible au bruit de fond des fichiers audio, pouvant induire en erreur l'analyste. Par exemple, concernant l'écart-type spectral, lorsqu'un signal décroît son intensité, la forme du spectre, en général, tend à s'aplatir augmentant, par conséquent, la valeur de l'écart-type spectral. Si ce signal possède du bruit de fond (qui en général présente une morphologie spectrale plutôt plate), cela n'est que plus vrai. Par exemple soit l'attaque percussive du premier canal, de la version quadriphonique, de *Bohor*, à 230.50 s (3 min 50.5 s). Le sonogramme et le spectre instantané sont représentés à la Figure 34. À cet instant le centroïde spectral est au tour de 2600 Hz et l'écart-type spectral au tour de 1500 Hz. Cela est facilement compréhensible, puisque le barycentre du spectre est plus bas, comme nous pouvons le voir dans cette figure.



Figure 34. *Bohor*, premier canal, sonogramme et spectre instantané à 3 min 50.5 sec.

Par contre, il suffit d'avancer un peu plus dans le fichier et avancer jusqu'à 238 s (3 min 58 s), où nous devrions avoir du silence, l'événement percussif s'étant complètement arrêté. Ce que nous constatons est que nous avons encore un bruit de fond, avec un spectre plus plat, comme nous le montre la Figure 35.

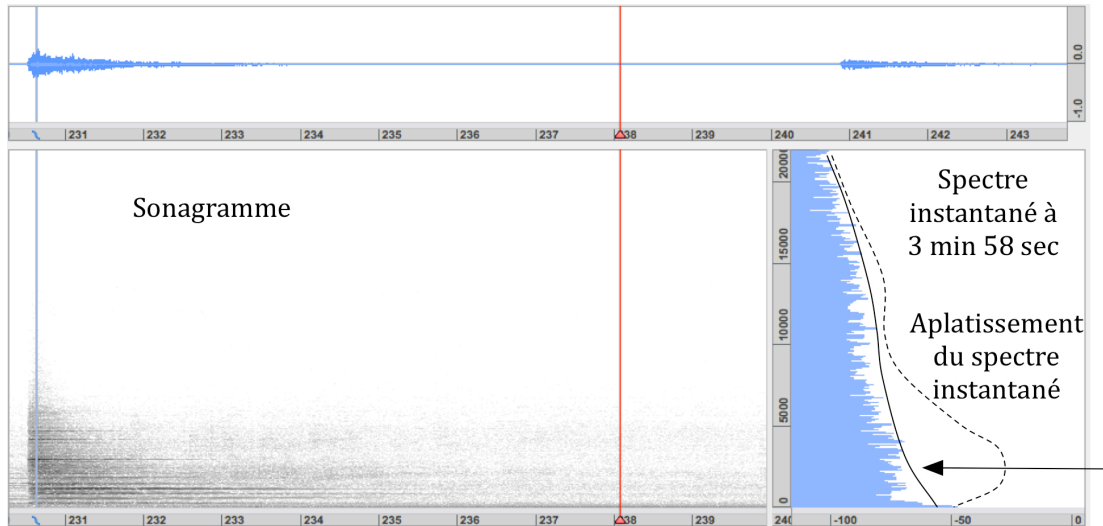


Figure 35. *Bohor*, premier canal, sonogramme et spectre instantané à 3 min 58 sec.

À 238 s (3 min 58 s), nous constatons une valeur d'écart-tipe spectrale au tour de 2800 Hz. Par l'aplatissement du spectre, non seulement l'écart-tipe spectral a augmenté, mais la valeur du centroïde spectral aussi. La valeur de l'écart-tipe spectral à ce moment est celle du bruit de fond. Ce phénomène se produit aussi avec des sons ayant un niveau très bas de bruit de fond. Cela montre le besoin d'un vrai modèle de l'écart-tipe perçu. Ce modèle n'existe pas encore à notre connaissance. Dans l'attente de ce modèle, aussi bien pour l'écart-tipe spectral que pour d'autres descripteurs, nous utilisons l'application d'une fonction de transfert, amenant la valeur de l'écart-tipe spectral à zéro à partir d'un niveau de seuil d'énergie du signal. Le niveau de seuil est, pour le moment, réglé « ad hoc » (Figure 36)¹⁸.

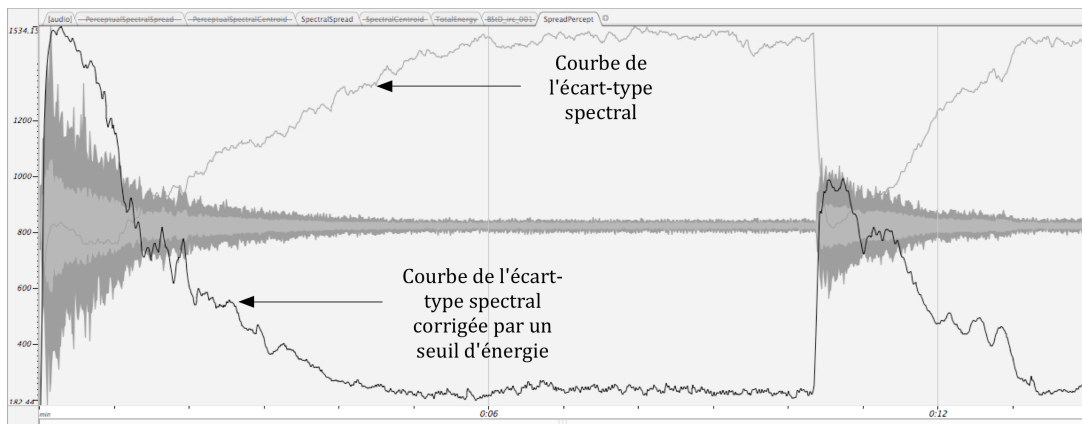


Figure 36. Correction de l'écart-tipe spectral, par rapport à un seuil d'énergie.

Un dernier point technique important est celui de la fenêtre temporelle utilisée pour l'analyse. Toutes ces méthodes utilisent à la base une analyse fondée sur la transformée de Fourier. Pour des analyses, avec une vocation globale et ayant l'intention de nous aider dans l'analyse aurale, il est aussi important de trouver la résolution temporelle convenable à nos objectifs. Par exemple, si nous sommes en train d'utiliser ces descripteurs pour travailler sur la forme, d'un point de vue global, des fenêtres de 2048 échantillons, soit 46 millisecondes, seraient sans doute trop petites, nous apportant trop d'information inutile, ou du bruit, pour l'échelle de notre étude. Des fenêtres d'une demi-seconde ou plus pourraient à ce moment convenir plus. Même si cela semble évident, l'échelle de l'observation, de l'analyse, devrait être en accord avec l'échelle de l'étude. Ces derniers commentaires nous montrent un point

¹⁸ Nous avons déjà brièvement expliqué ce point technique dans [Malt, Jourdan 2009], mais nous le développerons dans des travaux futures.

méthodologique qu'il serait utile de développer postérieurement, qu'est une recherche sur la granularité de ces représentations, de la « taille » moyenne de la fenêtre d'analyse, en fonction de notre perception et de notre mémoire.

Ces points techniques nous montrent que les réglages des analyses doivent être cohérentes avec nos objectifs, et que les musicologues qui voudraient utiliser et s'appropriier de ces outils, ne pourront pas faire l'économie d'une étude approfondie de leur fonctionnement.

Nous n'avons que survolé ces quatre pièces de Xenakis. Notre objectif, comme nous l'avions signalé était de tester une méthodologie d'analyse, fondée sur l'utilisation de descripteurs audio et de tester l'utilisation de la représentation de la masse sonore, que nous avions posée.

En ce qui nous concerne, nous pouvons dire que cette méthodologie, en plus de s'accorder avec des analyses préalables et reconnues, a aussi pu les développer en nous permettant de comprendre certains éléments pour inférer des articulations de la forme musicale. Les caractéristiques qui nous ont aidés ont été principalement les trajectoires des descripteurs et leurs discontinuités. Notamment en ce qui concerne la représentation de la masse sonore, la BStD, elle a permis de faire des rapprochements et des comparaisons. Des ces comparaisons nous avons pu inférer la similarité formelle entre *Diamorphoses* et *Orient-Occident*, nous amenons aussi au fait de l'utilisation par Xenakis de sections finales fondées sur la stabilité d'un paramètre sonore. Ce dont nous avons nommé des « pédales statiquement statiques », et qui se sont montrées (dans les cas de *Diamorphoses*, *Orient-Occident* et *Bohor*) de structure bipartite. Pour *Concret PH*, peut-être une des pièces les plus énigmatiques de Xenakis, de notre point de vue, nous avons aussi pu inférer une structure fondée sur la densité, la texture et le mouvement interne des grains.

Les œuvres électroacoustiques de Xenakis sont autant déconcertantes et dérangeantes qu'elles interpellent l'auditeur et l'analyste. Il n'est pas aisé de comprendre comment fonctionnait son écoute, qui, disons-le a maintes fois était critique, comme les « insuffisances d'oreille de Xenakis... » signalées par Tristan Murail [Murail 1989, 150].

Le fait de proposer une organisation formelle fondée sur l'étude de caractéristiques sonores nous permet aussi de lancer l'hypothèse que ces caractéristiques, que leurs conduites, les directions qu'elles articulent, pourraient être des indices nous permettant de comprendre l'écoute de Xenakis, confirmant qu'elle était effectivement une écoute globale, fondée sur une perception particulière de la masse sonore, des textures, des densités d'événements et des timbres, plus que sur des caractéristiques locales. De ces considérations, nous ne pouvons que convenir avec Makis Solomos [Solomos, 2009a] et [Solomos, 2010] sur le fait que, dans le contexte plus général de la musique de Xenakis, la forme s'explique mieux en termes d'articulation et de mise en relation d'énergies. La signification musicale sera construite à partir de sensations primitives de tension/détente, contraste/continuité, immobilité/mouvement, etc. de ces différentes caractéristiques sonores mises en rapport.

9. BIBLIOGRAPHIE

- Bregman, Albert S. (1990), *Auditory Scene Analysis*, MIT Press, Cambridge, MA.
- Chiasson, Frederic (2007), « L'orchestration selon Koechlin : recherche d'un point de rencontre entre la pratique musicale et les sciences cognitives », in *Actes du Colloque international « Composer au XXIème Siècle – Processus et Philosophies »*, OICM, Montréal (Québec) Canada, 28 février – 3 mars 2007.
- Couprie, Pierre (2007), « Une analyse détaillée de *Bohor* (1962) », *Definitive Proceedings of the International Symposium Iannis Xenakis* (Athènes, mai 2005), <http://cicm.mshparisnord.org/ColloqueXenakis/papers/Couprie.pdf>
- Di Scipio, Agostino (1997), « The problem of 2nd-order sonorities in Xenakis' electroacoustic music », in *Organised Sound*, vol. 2, n° 3, p. 165–78.
- Di Scipio, Agostino (1998), « Compositional Models in Xenakis's Electroacoustic Music », *Perspectives of New Music*, vol. 36, n°2, p. 201–243.

- Fraassen, Bas C. Van (2008), *Scientific Representation: Paradoxes of Perspective*, Oxford University Press,.
- Gibson, Benoit (2012), *Bohor* (1962) de Iannis Xenakis, http://dspace.uevora.pt/rdpc/bitstream/10174/8831/1/bohor_analyse.pdf
- Harley, James (2002), « The electroacoustic music of Iannis Xenakis. », in *Computer Music Journal* vol. 26, n°1, p. 33–57.
- Henning, Lohner (1986), « The UPIC System: A User's Report. » *Computer Music Journal* vol. 10, n° 4, MIT Press, p. 42–49.
- Malt, Mikhail, Emmanuel Jourdan (2009), « La « BSTD » – Une représentation graphique de la brillance et de l'écart type spectral, comme possible représentation de l'évolution du timbre sonore », dans le cadre de l'édition des textes du colloque international *L'ANALYSE MUSICALE AUJOURD'HUI, Crise ou (r)évolution ?*, Université de Strasbourg/SFAM, 19-21 novembre 2009, 19/11/2009. Edition de Xavier Hasher et Mondher Ayari, (2009), en cours de réalisation.
- Murail, Tristan (1989), « Questions de cible », in *Entretiens* n° 8, Paris, p. 147-172.
- Peeters, Geoffroy (2003), *A Large Set of Audio Features for Sound Description*, rapport technique, Ircam, Paris, France.
- Pressnitzer, Daniel, Stephen McAdams, Suzanne Winsberg, Joshua Fineberg (2000), « Perception of musical tension for non-tonal orchestral timbres and its relation to psychoacoustic roughness », in *Perception and Psychophysics*, vol. 62, n° 1, p. 66–80. http://audition.ens.fr/dp/pdfs/pressnitzer-2000-orchestral_timbres.pdf.
- Schaeffer, Pierre (1966), *Traite des objets musicaux*, Paris, Le Seuil, Essais interdisciplines,.
- Siedenburg, Kai (2009), *An Exploration of Real-Time Visualizations of Musical Timbre*, Report. <http://cnmat.berkeley.edu/system/files/attachments/Paper.pdf>
- Solomos, Makis (1994), « Les trois sonorités xenakiennes », in *Circuit : musiques contemporaines*, vol. 5, n° 2, p. 21–40.
- Solomos, Makis (2001), « The Unity of Xenakis's Instrumental and Electroacoustic Music: The Case for “Brownian Movements” », in *Perspectives of New Music*, vol. 39, n° 1, p. 244–254.
- Solomos, Makis (2005), « Le Diatope et La Légende d'Eer », www.iannis-xenakis.org/fxe/actus/Solom3.pdf.
- Solomos, Makis (2009a), « *Orient-Occident*. From the film version to the concert version », in *Iannis Xenakis : Das elektroakustische Werk. Internationales Symposium Musikwissenschaftliches Institut der Universität zu Köln*, sous la direction de Ralph Paland, Christoph von Blumröder, Vienne, Verlag der Apfel, p. 118–131.
- Solomos, Makis (2009b), « Pour une approche analytique des phénomènes musicaux énergétiques », Communication au colloque *L'analyse musicale aujourd'hui : Crise ou (r)évolution ?*, Université de Strasbourg et la SFAM, Strasbourg, France.
- Solomos, Makis (2010), « La musique comme énergie », *Programme du festival Musica, Festival des musiques d'aujourd'hui*, Strasbourg, p. 36 - 38.
- Schnell, Norbert, A. Röbel, D. Schwarz, G. Peeters, R. Borghesi (2009), « MuBu & Friends - Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP », in *International Computer Music Conference (ICMC)*, Montreal. <http://articles.ircam.fr/textes/Schnell09a/>
- Squibbs, Ronald (1996), « Images of Sound in Xenakis's Mycenae Alpha », in *Journées d'informatique musicale*, L'Ile de Tatihou, Normandy, France.
- Sterken, Sven (2004), *Iannis Xenakis, Ingénieur et Architecte. Une analyse thématique de l'œuvre, suivie d'un inventaire critique de la collaboration avec Le Corbusier, des projets architecturaux et des installations réalisées dans le domaine du multimédia*, Thèse de doctorat déposée pour l'obtention du degré de docteur en Sciences Appliquées, Section Architecture, Université de Gent, 2003-2004, Directeur de Thèse : Prof. Dr. Bart Verschaffel.
- Tiffon, Vincent (2006), « La représentation sonographique est-elle une aide pour l'analyse perceptive de la musique électroacoustique? », *Revue Liens, Editions*

- Musiques&recherches*, p. 3-15, <http://www.musiques-recherches.be/images/stories/documents/LienAnalyse2011.pdf> (21/08/2013)
- Valsamakis, Nikolas (2009), *Aesthetics and Techniques in the Electroacoustic Music of Iannis Xenakis*, Pella Publishing Company. <http://hdl.handle.net/10066/5777>.
- Vassilakis, Pantelis N., Auditory Roughness, <http://musicalgorithms.ewu.edu/learnmoesra/moreroughness.html>.
- Vassilakis, Pantelis N. (2001), *Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance*, Thèse de doctorat en musicologie systématique, Los Angeles, University of California, Los Angeles.
- Vassilakis, Pantelis N. (2005), « Auditory roughness as a means of musical expression », in *Selected Reports in Ethnomusicology* (Perspectives in Systematic Musicology), vol. 12, p. 119-144.
- Xenakis, Iannis (1963), *Musiques formelles*, Stock Musique, 261 p.
- Xenakis, Iannis (1987), « *Mycenae-Alpha* », in *Perspectives of New Music*, vol. 25, n° 1/2, p. 12-15.

10. DISCOGRAPHIE

- Xenakis, Iannis (1997a), « *Diamorphoses* (1957) », *Xenakis: Electronic Music*, Electronic Music Foundation, EMF CD 003, 1997.
- Xenakis, Iannis (1997b), « *Concret PH* (1958) », *Xenakis: Electronic Music*, Electronic Music Foundation, EMF CD 003, 1997.
- Xenakis, Iannis (1997c), « *Orient-Occident* (1960) », *Xenakis: Electronic Music*, Electronic Music Foundation, EMF CD 003, 1997.
- Xenakis, Iannis (1997d), « *Bohor* (1962) », *Xenakis: Electronic Music*, Electronic Music Foundation, EMF, CD 003, 1997.
- Shaeffer, Pierre (1990), *L'œuvre musicale*, INA-GRM.