Tension et rugosité dans la musique non-tonale.

Daniel Pressnitzer, IRCAM, 1 Pl. Stravinsky, F-75004 Paris, pressnitzer@ircam.fr
Stephen McAdams, IRCAM et Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, EPHE, 28 rue Serpente, F-75006 Paris. smc@ircam.fr
Suzanne Winsberg, IRCAM, winsberg@ircam.fr
Joshua Fineberg, IRCAM, fineberg@ircam.fr

Résumé: *L'objectif de cette étude est de déterminer si la notion musicale de tension et détente peut rester pertinente dans un cadre non-tonal, et dans quelle mesure elle peut être expliquée et modélisée par la rugosité psychoacoustique. Huit accords de la pièce "Streamlines" (J. Fineberg, 1995) ont été enregistrés avec des instruments acoustiques, de manière à constituer des timbres orchestraux fusionnés. Une première expérience montre qu'une hiérarchie est clairement perçue par les sujets, mais que les échelles de tension et de rugosité sont significativement distinctes. Les différences seraient dûes à des spécificités contenues dans certains des timbres. Une deuxième expérience sur les mêmes timbres orchestraux remixés permet de confirmer cette hypothèse, et indique cette fois deux échelles confondues. Un modèle basé sur les résultats de Hutchinson et Knopoff (1978) fournit des prévisions satisfaisantes. La différence observée entre les deux conditions expérimentales permet de discuter du champ d'application potentiel d'un tel modèle.*

1 Introduction

L'étude des mouvements de tensions et détentes au cours du temps constitue la base des analyses harmoniques classiques. Par exemple, lors d'une cadence, le septième degré d'une gamme majeure est considéré comme instable, dissonant, et appelle une résolution sur la tonique, stable. Cet instabilité constitue une tension, sa résolution une détente. Cette notion est considérée comme fondamentale par nombre de théoriciens de la musique tonale [Schenker, 1935]. [Lerdahl and Jackendoff, 1983] posent l'existence d'un tel réseau de tension et détente comme axiome élémentaire d'un des composants de leur théorie générative de la musique : la réduction prolongationnelle. Les éléments faibles pour leur hiérarchie sont source de tension, résolue par le passage à un élément de valeur hiérarchique plus forte. [Bigand, 1993] a démontré que ces mouvements de tension étaient perçus à la fois par des sujets musiciens et non-musiciens, pour des mélodies et pour des accords, soulignant donc leur importance dans l'écoute musicale. Ces études révèlent aussi l'influence de l'acculturation implicite de l'auditeur : ainsi, un compositeur de musique tonale peut-il s'appuyer efficacement sur un ensemble de règles désormais connues pour faire ressentir ces mouvements.
Le problème est sensiblement différent pour un compositeur ne désirant pas faire référence à la tonalité. Prenons l'exemple de la musique dite "spectrale". Le point commun des compositeurs regroupés (à tort ou à raison) sous cette étiquette est la volonté de bâtir un système harmonique basé sur la structure du son lui-même. Le défi d'une telle approche est de donner à l'auditeur des points de référence naturellement compréhensibles, car contenus dans le matériau, tout en permettant d'exploiter les possibilités nouvelles offertes par l'informatique musicale dans le domaine de la synthèse ou de l'assistance à la composition. La genèse de la séquence d'accords choisie pour constituer la matière de l'étude est à ce titre exemplaire. La partition de l'extrait de "Streamlines" (J. Fineberg) réorchestré par le compositeur pour le besoin de l'étude est donnée en figure 1. Un réservoir de matériaux a d'abord été constitué par analyses acoustiques de sons de contrebasse dans des modes de jeu inhabituels. Les accords ont ensuite été choisis par le compositeur et agencés de manière à suivre un contour précis de tensions qui était requis pour des raisons musicales. Ces accords n'ont aucune fonction tonale simple. Un moyen inédit de les hiérarchiser au niveau de la tension est donc nécessaire. Le critère retenu pour classer ces accords est de leur attribuer une valeur basée sur un calcul du plus grand commun diviseur des notes fondamentales qui les constituent. Plus cette valeur est élevée,

FIG. 1 – *Streamlines (J. Fineberg). Partition de l'extrait utilisé. Les accords indiqués au-dessus de la partition ont été enregistrés isolement, sans rythme, avec une même dynamique, et constituent les huits timbres orchestraux soumis à l'expérimentation.*

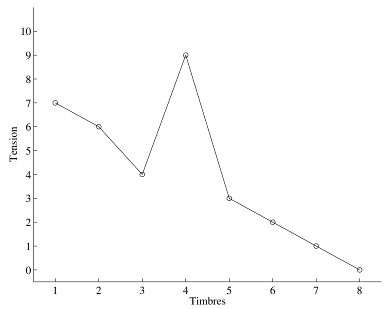

FIG. 2 – *La tension calculée par l'algorithme original.*

plus l'accord est qualitativement considéré comme porteur de tension (figure 2, la morphologie seule de la courbe a été utilisée par le compositeur plutôt que la valeur numérique attribuée à chacun des timbres.). Plusieurs questions sont alors légitimes : la notion de tension reste-t-elle pertinente pour de tels accords ne rentrant pas dans la théorie harmonique classique? Dans la pièce réelle, le mouvement est souligné par le rythme et les dynamiques et le contour de tension est effectivement clairement perceptible. Toutefois, l'indice utilisé pour le calcul précompositionnel est purement spectral : pour être cohérent, la tension doit être fondée sur la seule base des différences de timbre entre ces accords. L'expérience présentée a pour premier but de démontrer l'existence claire d'une hiérarchie de tension non-tonale ne s'appuyant ni sur le rythme, ni sur les différences de dynamique.

Le deuxième objectif de l'expérimentation est d'explorer le lien potentiel entre la grandeur psychoacoustique appelée "rugosité" et cette tension non-tonale. La rugosité psychoacoustique est un attribut perceptif du son qui a été introduit par [Helmholtz, 1954]. Elle est liée à la perception de fluctuations d'amplitude rapides, d'une fréquence de modulation située entre 20 et 200 Hz. Deux sons proches en fréquence entendus simulanément vont donner lieu à des battements qui peuvent être rugueux. Une seconde mineure jouée dans le registre medium ou grave du piano en est un exemple. La rugosité a fait l'objet de nombreuses études expérimentales, qui se sont attachées à quantifier sa perception et à proposer des mécanismes auditifs permettant de la comprendre.[Plomp and Levelt, 1965] ont permis d'établir le lien entre rugosité et bande critique grâce à une étude portant sur des paires de sinus dont le registre et l'écart en fréquence varient [1]. [Terhardt, 1974] a adopté un point de vue légèrement différent en introduisant l'étude de sinus modulés en amplitude, soulignant ainsi le caractère temporel du phénomène. D'autres études plus récentes avec des sons plus complexes incluant des

1. La bande critique ([Fletcher, 1940]) définit la résolution fréquentielle de l'oreille, et découle principalement de la mécanique de la membrane basilaire.

bruits ([Aures, 1985]) confirment le lien étroit entre la perception de la rugosité et certaines caractéristiques du système auditif, comme l'analyse fréquentielle effectuée par la membrane basilaire et la caractéristique passe-bas de la perception temporelle de la modulation d'amplitude.
Toutes ces études tendent donc à démontrer l'existence d'un attribut sonore clairement défini lié à des mécanismes auditifs périphériques. Si un lien peut être établi entre cet attribut quantifiable et la notion fondamentale de tension, cela laisse entrevoir la possibilité de réfléchir à un modèle de calcul prévisionnel pouvant présenter un réel intérêt musical.

2 Partie expérimentale

La partie expérimentale à pour objet d'établir, si cela est pertinent, une échelle de tension et une échelle de rugosité pour les accords présentés en figure 1. De manière à éviter un jugement direct difficile à définir sans introduire de quelconques adjectifs induisant une préférence esthétique, une procédure de comparaison simple entre accords arrangés par paires est adoptée. Une première expérience est menée avec les accords enregistrés en prise directe par des instruments acoustiques, en "situation de concert". Cet enregistrement contient toutes les spécificités liées à l'interprétation, naturelles dans une telle situation. Une deuxième expérience, selon le même protocole, est réalisée avec une version remixée des accords, et représente une version "discographique" conforme aux vœux du compositeur. Les accords diffèrent entre eux principalement par leurs caractéristiques spectrales, et non par leur dynamique ou une quelconque fonction tonale : nous préférons donc nous y référer comme à des *timbres orchestraux*.

2.1 Expérience "prise directe"

2.1.1 Méthode

Stimuli : Huit accords extraits de la pièce "Streamlines" (Joshua Fineberg, 1995) ont été interprétés par des instrumentistes de l'ensemble "Itinéraire" (deux flûtes, une clarinette, un alto et un violon). Les huits accords ont été joués isolément, sous la direction du compositeur, avec une même durée, selon la même nuance (dal niente - mezzo forte - dal niente). Ils constituent les huits timbres orchestraux dont nous parlerons désormais.

Appareillage : L'enregistrement a été effectué dans la salle de concert de l'IRCAM, avec un couple de microphones ORTF Schoeps placé à 5m des instrumentistes. La prise stéréo a été reportée numériquement sur disque dur à 44.1 kHz de fréquence d'échantillonage et 16 bits de résolution dynamique. Lors de l'expérience, le sujet était enfermé seul dans une cabine insonorisée, et répondait au moyen du programme PsiExp de Benett Smith, implanté sur une station NeXT. Les stimuli étaient diffusés sur un casque ouvert AKG K1000. Le niveau d'écoute mesuré était autour de 80 dBA, modifiable sur demande.

Sujets : Le groupe expérimental était constitué de 29 sujets, et ne comportait pas les auteurs. Les sujets avaient de 17 à 45 ans, avec une moyenne de 26 ans, et provenaient d'horizons divers : certains travaillant à l'IRCAM, d'autres recrutés par base de données. Ils étaient payés pour leur participation à l'expérience. Le groupe comporte 9 musiciens professionels ou semi professionels (compositeurs et instrumentistes), 8 musiciens amateurs (instrumentistes occasionnels) et 12 non-musiciens.

Protocole : L'expérience comporte deux parties, l'une portant sur la tension l'autre sur la rugosité. Les huits timbres sont arrangés en 56 paires (toutes les combinaisons de timbres différents dans un ordre comme dans l'autre) qui sont présentées en ordre aléatoire. La tâche consiste à faire un jugement comparatif à choix forcé entre les deux éléments de la paire. La question figurant sur la consigne écrite est : *Entre les deux sons de cette paire, percevez-vous une évolution du type tension-détente ou au contraire du type détente-tension?*. Il est précisé que les sujets peuvent employer les critères qui leur semblent appropriés pour définir les mots "tension" et "détente". Après cinq essais de familiarisation, cette première partie dure environ 30 mn. La deuxième partie consiste à focaliser l'attention du sujet sur un attribut perceptif particulier, la rugosité. Cet attribut est introduit par un exemple, un sinus de 1000 Hz modulé à 70 Hz, dont les sujets peuvent faire varier eux-mêmes la profondeur de modulation de 0 % à 100 %. La rugosité de l'exemple passe ainsi de 0 à 1 asper (selon [Zwicker and Fastl, 1990]). La tâche consiste ensuite, de facon similaire à la première partie, à faire un jugement comparatif "plus rugueux-moins rugueux" ou "moins rugueux-plus rugueux" pour toutes les paires.

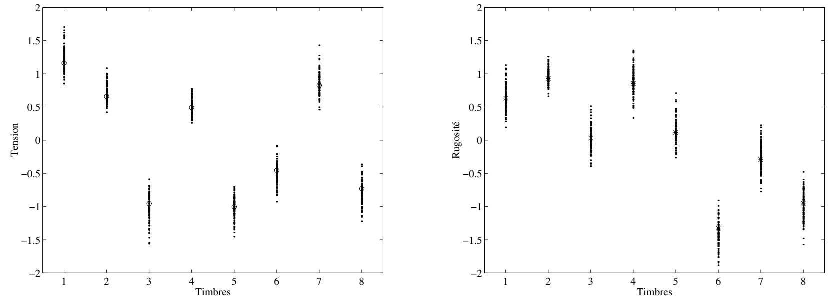

FIG. 3 – *Échelles de tension (à droite) et de rugosité (à gauche) pour les huits timbres obtenues par analyse BTL, et résultats des 100 bootstrap.*

2.1.2 Hiérarchie de tension et de rugosité

De manière à pouvoir comparer les résultats de l'expérience et les prévisions compositionnelles, nous devons arranger les huits timbres selon une échelle linéaire. L'expérience fournit pour chaque sujet 56 jugements de supériorité de tension ou de rugosité, correspondant aux comparaisons de toutes les paires de timbres dans un ordre comme dans l'autre. La méthode de Bradley-Terry-Luce (BTL, [David, 1988]) a été adoptée pour définir cette échelle. Nous faisons l'hypothèse que chacun des timbres T_i possède une valeur "vraie" de tension θ_i, correspondant à l'échelle recherchée. Lors d'une comparaison entre deux timbres i et j, le sujet ne compare pas directement les valeurs θ_i et θ_j recherchées mais plutôt les variables aléatoires associées t_i et t_j, centrées sur θ_i et θ_j. Dans l'hypothèse BTL, la variable aléatoire suit une loi en sécante hyperbolique. Ainsi, la probabilité que θ_i soit jugé supérieur à θ_j est donnée par l'équation (1). Les probabilités π_{ij} sont estimées par la proportion de jugements obtenus p_{ij}, et en inversant (1) on obtient les estimés t_i des θ_i.

$$\pi_{ij} = H(\theta_i - \theta_j) = 1/2(1 + tanh(1/2(\theta_i - \theta_j))) \tag{1}$$

Ainsi, nous disposons pour chaque timbre d'une valeur de tension et d'une valeur de rugosité. Pour vérifier si les valeurs diffèrent significativement entre deux stimuli, il faut examiner si leur écart est supérieur à l'écart-type des résultats. Ceci est justement ce que fait une analyse de variance. Toutefois, une analyse de variance présume une erreur normale, ce que nous ne voulons pas affirmer à priori. La technique du bootstrap permet d'estimer la stabilité d'un paramètre comme l'écart-type en utilisant une distribution empirique, par opposition à l'hypothèse de distribution normale [Efron, 1981]. La distribution empirique est obtenue par tirage avec remise de l'échantillon expérimental de sujets réel. Nous avons effectué 100 réplications de l'expérience par bootstrap, en tirant au sort avec remise 29 sujets de notre groupe expérimental, ce qui constitue une taille adéquate pour estimer l'écart-type [Efron and Tibshirani, 1993]. Les résultats de ces 100 réplications traités eux aussi par la méthode BTL sont représentés figure 3. La valeur BTL obtenue avec le groupe de sujets réel est indiquée par une étoile.

Il apparaît que les stimuli présentent des différences significatives selon les échelles considérées. L'examen des nuages de points indique que l'erreur expérimentale n'est pas normale, justifiant la méthode d'analyse employée. La figure 4 permet une comparaison des deux échelles obtenues, en présentant les valeurs BTL réellement obtenues, reliées entre elles pour faciliter la lecture, et les erreurs standard estimées par bootstrap. Il existe des différences significatives entre les deux échelles. De plus, les courbes obtenues ne sont pas en accord avec les prévisions compositionnelles.

2.1.3 Différence entre les échelles

Grâce à l'analyse par bootstrap, nous pouvons affirmer qu'il existe une échelle significative de tension non-tonale entre les différents timbres étudiés. La rugosité est elle aussi une grandeur pertinente pour caractériser ces timbres, au vu de la stabilité des résultats obtenus. Toutefois, on peut remarquer que les deux échelles diffèrent significativement en plusieurs endroits. Ceci indique qu'il existe des

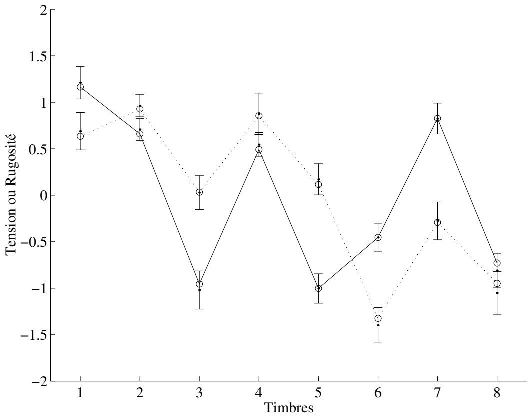

FIG. 4 – *Comparaison des échelles de tension (trait plein) et de rugosité (pointillés), avec écarts type estimés par bootstrap*

facteurs autres que la rugosité qui influent sur les jugements de tension des sujets. Nous faisons l'hypothèse que parmi ces facteurs, il en est qui sont dûs à des spécificités d'interprétation, liées à la situation de concert. Un premier type de spécificités est une "spécificité de surface". Pour les timbres T_8, T_6 et T_7 les flûtes doivent jouer dans un registre aigü, et ceci à mezzo piano. Cette nuance est extrêment difficile à obtenir dans un tel registre, et la note de flûte ressort des timbres. Un deuxième type de spécificité est une "spécificité tonale". Le timbre T_3 contient un accord presque parfait majeur : une quinte juste (si triple dièse - fa quart de dièse entre les deux flûtes), et une tierce presque majeure (fa noté en si bémol donc ré dièse pour la clarinette). Ces intervalles auraient dû disparaître perceptivement avec l'impression de fusion générale voulue par le compositeur, mais du fait d'un deséquilibre entre les instruments ils deviennent saillant. Nous voudrions vérifier cette hypothèse, et préciser le rapport entre tension et rugosité pour ce qui a été écrit et pensé par le compositeur : une deuxième expérience est donc mise en place, où les spécificités sont diminuées grâce à un remixage des accords destiné à favoriser davantage la fusion. Cette manipulation a été effectuée sous la supervision du compositeur, de manière à rester musicalement réaliste, et pourrait être le résultat discographique de l'enregistrement de concert.

2.2 Expérience "remixage"

2.2.1 Méthode

Stimuli: Les mêmes accords que ceux de la condition "prise directe" sont utilisés. Nous avions utilisé une prise stéréophonique de ces accords, mais nous disposions aussi d'un enregistrement multipistes avec chaque instrument sur une piste. Les volumes des différents instruments ont été homogénéisés de manière à obtenir une meilleure fusion.

Appareillage : La prise de son a été réalisée avec 5 microphones KM 140 Neuman, placés chacun à 50 cm d'un instrumentiste. Le mixage a été réalisé sur une table Neve série V à partir d'un lecteur multipiste 33.24 A Sony vers un DAT 7050 Sony. Une réverbération artificielle Lexicon a été rajoutée pour rendre un effet de salle semblable à celui présent sur l'enregistrement du couple.

Sujets : 10 sujets âgés de 18 à 35 ans ont passé cette expérience, avec une moyenne d'âge de 27 ans. Ces sujets n'avaient pas passé la condition "prise directe", mais ont été recrutés de la même manière. Le groupe compte 4 non-musiciens, 4 musiciens amateurs et 2 musiciens professionnels ou semi-professionels. Ils étaient rémunérés pour passer l'expérience.

Protocole expérimental : le protocole a été maintenu strictement identique entre les deux expériences.

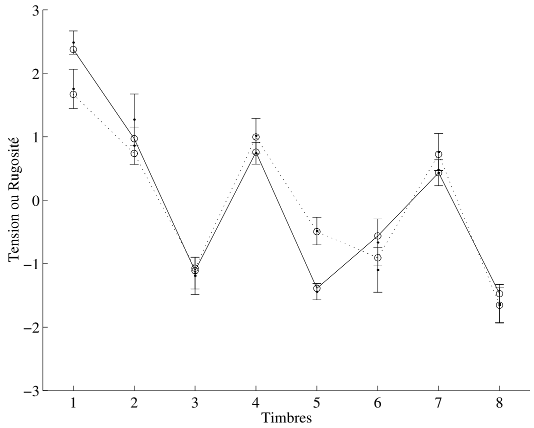

FIG. 5 – *Comparaison des échelles de tension (trait plein) et de rugosité (pointillés), avec erreurs standard estimées par bootstrap, pour les timbres de l'expérience "remixage".*

2.2.2 Hiérarchie de tension et de rugosité

La figure 5 représente les résultats des analyses BTL pour les nouvelles données, ainsi que l'écart-type estimé par bootstrap. Elle est l'équivalent de la figure 3 avec les timbres rééquilibrés. Les courbes de tension et de rugosité évoquent la hiérarchie obtenue précédemment, ce qui est dû à la très grande similitude des stimuli employés dans les deux conditions. Les résultats obtenus sont encore une fois très stables, indiquant la pertinence des dimensions jugées. Les différences entre les stimuli sont hautement significatives. Nous confirmons ainsi certains résultats, et ce avec un groupe de sujets indépendant. Le résultat nouveau est que cette fois la rugosité perçue est très fortement corrélée à la tension musicale (le coefficient de corrélation entre les deux courbes est $r = 0.94$, intervalle de confiance à 95 % estimé par boostrap et percentile : $r \in [0.87\ 0.96], \alpha = 0.05$). Cette corrélation a été obtenue après rééquilibrage des niveaux d'enregistrement de nos stimuli, ce qui a eu pour effet d'augmenter leur degré de fusion perceptive. Toutefois, la hiérarchie obtenue diffère de celle prévue par l'algorithme basé sur les fondamentales virtuelles présenté en introduction de cette étude (figure 2). Si la première partie de la courbe a été bien pressentie, la deuxième partie est en désaccord avec les prévisions, notamment pour le timbre 7 prévu comme "détendu". Pourtant ce timbre est bien porteur de tension, comme l'indiquent les jugements. Or il est aussi perçu comme rugueux : la rugosité semble donc être un meilleur prédicteur de la tension évoquée. Un modéle de rugosité pourrait donc apporter une aide appréciable pour l'interprétation et l'établissement prévisionnel de l'échelle de tension. C'est un tel modèle que nous allons examiner, à la lumière de nos résultats expérimentaux.

3 Modélisation

3.1 Présentation du modèle de calcul de rugosité

Le modèle envisagé découle directement de l'étude de [Hutchinson and Knopoff, 1978], les base de cette étude étant les résultats expérimentaux de [Plomp and Levelt, 1965] cités dans l'introduction. Le modèle prend comme données initiales les fréquences F_i et amplitudes A_i des différents composantes spectrales du son à étudier. Ces fréquences et amplitudes sont à rapprocher à l'idée de densité spectrale moyenne, mais les amplitudes restent hautement indicatives. Chaque composant est comparé à tous les autres. Pour chaque couple de composants, un nombre adimensionnel g correspondant à la courbe proposée par [Plomp and Levelt, 1965] est calculé en fonction du rapport entre leur écart fréquentiel et la largeur calculée de la bande critique à la fréquence moyenne considérée. La bande critique est calculée selon la formule proposée par [Hutchinson and Knopoff, 1978], le nombre g grâce à une formule de [Bigand et al., 1996]. Deux composants éloignés de plus de 120 % de la

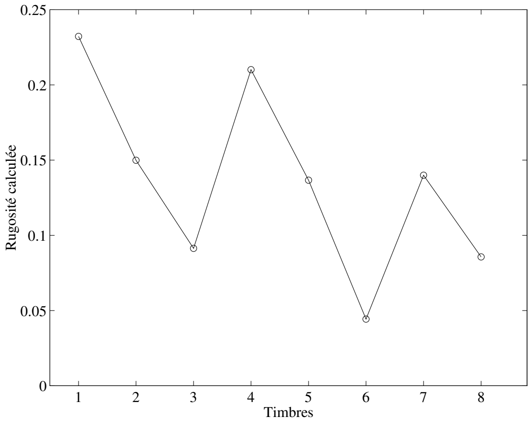

FIG. 6 – *Calcul de rugosité pour les huits accords expérimentaux, en considérant 4 harmoniques pour chaque fondamentale.*

bande critique ont un g égal à zéro, deux autres éloignés de 25 % de la bande critique ont un g de 1. Les différents g partiels sont ensuite sommés, en normalisant quant aux amplitudes. Le résultats du calcul est donné par (2).

$$D = \frac{\sum_{i=1 j=1}^{n} a_i a_j g}{\sum_{i=1}^{n} a_i^2} \tag{2}$$

Ce modèle a été implémenté dans l'environnement d'aide à la composition *Patchwork*, grâce à G. Assayag. Il comporte un module de calcul proprement dit, qui peut recevoir un ensemble de notes ou de fréquences (provenant d'une partition, par exemple). Il comporte aussi un générateur de composantes fréquentielles à partir de fondamentales, car la forme du "spectre" utilisé (nombre d'harmoniques et amplitudes respectives) a une influence déterminante sur le résultats des calculs. En effet, les battements entre harmoniques constituent une source essentielle de rugosité. Si l'écart entre les fondamentales dépasse une bande critique, les battements entre harmoniques sont même la seule source de rugosité. Il importe donc d'émettre une hypothèse sur la forme réelle des spectres qui vont être générés au moment de la réalisation sonore de la partition.

3.2 Évaluation du modèle avec les stimuli utilisés

Ce modèle très simple est destiné à une phase précompositionnelle, et les données choisies pour l'alimenter sont donc les fréquences fondamentales constituants les accords écrits par le compositeur. Ces fréquences fondamentales ont été associées à un spectre d'amplitude. Compte tenu du registre et des instrument employés, des "spectres" harmoniques de 4 partiels avec une décroissance d'amplitude en $1/n^2$ sont employés. Ceci est bien entendu une grande simplification, mais reste plausible. Les résultats du calcul par le modèle sont présentés figure 6. Il existe un bon ajustement entre les résultats du modèle et les jugements expérimentaux de rugosité de l'expérience avec les timbres remixés (figure 5). Le coefficient de corrélation des deux échelles est de $r = 0.88$, $r \in [0.81\ 0.94], \alpha = 0.05$). Le modèle permet donc une bonne prédiction de la rugosité perçue, et traduit notament le fait que le timbre T_7 ait été jugé rugueux. Du fait de la similitude entre les deux échelles, le modèle donne donc aussi une bonne prévision de la tension musicale jugée ($r = 0.79$, $r \in [0.69\ 0.85], \alpha = 0.05$)). Ceci est un résultat intéressant, mais qu'il convient de tempérer du fait des limites connues de ce modèle.

3.3 Limites du modèle

Ce modèle, destiné à un calcul exploratoire, comporte bien des faiblesses théoriques. Tout d'abord, la formule utilisée pour calculer la bande critique a depuis été révisée [Moore and Glasberg, 1983]. Le modèle ne reproduit les expériences de [Plomp and Levelt, 1965] qu'avec une formule que l'on sait donc désormais incorrecte. Plus grave, la prise en compte des amplitudes n'est fondée sur aucune étude expérimentale, mais sur une conjecture de sommation linéaire de rugosités partielles. Les seules études expérimentales prises en compte sont basées sur l'interaction de 2 sons purs, à partir desquels on extrapole à n sons purs. On sait désormais que l'influence du niveau est plus complexe, comme l'avaient d'ailleurs déjà remarqué [Kameoka and Kuriyagawa, 1969]. Cette influence, pour un son complexe, ne peut être dissociée des problèmes de phase et de cohérence entre les signaux en interaction. Il convient donc de discuter dans quelle mesure les résultats d'un tel modèle peuvent s'avérer dignes d'intérêt.

4 Discussion

Cette étude est le fruit d'une collaboration avec un compositeur, et ceci nous apparaît important à souligner. Les stimuli soumis à l'expérimentation ont été tirés d'une œuvre existante. Ils ont été enregistrés avec des instruments acoustiques en situation de concert, puis remixés dans une situation proche de l'édition discographique. Le compromis inévitable entre le contrôle rigoureux de tous les paramètres mis en jeu et la validité écologique de l'expérience est ici consciemment bien clair : nous avons opté pour une situation musicalement réaliste. La question que nous abordons, la hiérarchie de tension des timbres, est motivée par les préoccupations rencontrées par le compositeur au moment de l'écriture de la pièce. Les conditions particulières de l'étude nous ont amené à des méthodes d'analyses statistiques non paramétriques, c'est à dire ne supposant pas les hypothèses de normalité ne pouvant être vérifiées que dans des situation expérimentales comportant un plus grand nombre de sujets. Après une étude expérimentale, nous sommes en mesure de préciser la grandeur psychoacoustique, déjà connue et étudiée par ailleurs, qui est sous-jacente au problème posé. Nous pouvons aussi préciser les limites d'une telle interprétation, ce qui est peut-être aussi important que l'interprétation elle-même.

Nous avons projeté les différences entre les timbres, présentés par paires, sur des échelles linéaires. La dimension de rugosité s'est révelée bien présente dans les stimuli, comme le montre la cohérence des réponses obtenues, et pertinente musicalement, du fait de sa corrélation presque parfaite avec la hiérarchie de tension dans la deuxième condition expérimentale. Mais la première expérience montre aussi que d'autres facteurs sont sans doute présents dans la perception de la tension musicale de timbres orchestraux. La sonie est sans doute influente, même si les instrumentistes ont joué à la même nuance. Nous n'avions alors pas, dans un premier temps, retouché à cette interprétation pour se placer en situation de concert. La hauteur globale percue de timbres peut aussi sans doute avoir une influence. Un facteur lié à la centroïde spectrale pourrait aider à expliquer la tension évoquées par les timbres T_6, T_7 et T_8 lors de la première expérience, en rendant compte de la saillance des notes de flûte aigües.

Nous avons aussi tenté de supprimer les effets dûs au contexte et à la succession temporelle des timbres : en effet, tous les ordres possibles de paires sont présentés aléatoirement. Or la perception s'inscrit dans le temps, et l'influence l'évolution de la représentation mentale de la forme musicale au cours du temps sont ici ignorées. Nous avons attribué à chacun de nos timbre *une* valeur de tension et de rugosité : or ces timbres possèdent un début et une fin, la valeur attribuée représente-t-elle une moyenne, un maximum, un instant précis ? N'est-elle pas susceptible de varier en fonction du contexte, et des processus d'écoute volontaire ? Les réponses à ces question sont sans doute positives, et restent à préciser.

Ceci précise donc ce que l'on peut attendre et ce que l'on ne peut pas attendre d'un modèle de rugosité comme celui présenté. Le modèle nous a permis de prévoir la rugosité jugée par les sujets dans le cas de l'expérience "remixage". Les timbres étaient alors homogènes au niveau des dynamiques et des registres. Ils possédaient une grande fusion perceptive. Ils ne présentaient pas non plus de fonction tonale simple susceptible d'influencer le jugements des sujets par une sorte de référence extérieure difficilement contrôlable. Si, par contre, comme cela était le cas avec l'expérience "prise directe", ces conditions d'homogénéité ne sont pas assurées, alors les prévisions du modèle deviennent non pertinentes. Dans tous les cas un tel modèle ne peut rendre compte de l'influence du contexte, de la mémorisation et des processus d'attention volontaire. De fait, ses imprécisions théoriques sont en quelque sorte masquées par les limites inhérentes à l'approche adoptée. Néanmoins, son utilisation, si elle reste consciente des limites des nombres étiquetés sur le matériau musical peut constituer une aide à la composition inspirée de principes perceptifs plutôt que de théories purement arbitraires.

5 Conclusion

Nous avons alors montré que des stimuli orchestraux complexes, différant principalement par leur timbre, sans fonction harmonique classique, peuvent évoquer des mouvements de tension et de détente. Ces mouvements sont perçus à la fois par des sujets musiciens ou non musiciens, et entretiennent un lien avec la rugosité des différents stimuli. Ce lien est d'autant plus étroit que diverses spécificités sont maintenues à l'arrière plan. Si le matériau est homogène, la rugosité psychoacoustique devient une grandeur remarquable : en effet, ce phénomène lié à des mécanismes auditifs périphériques est une caractéristique de surface corrélée avec une notion musicale fondamentale. Son calcul prévisionnel par un modèle simple est alors possible.

Ceci peut présenter un réel intérêt musical. En effet, la perception de degrés de tension distincts pour nos timbres est sans doute plus fragile que dans un contexte tonal, elle peut notamment être modulée par une interprétation particulière, mais elle est démontrée. Un geste musical peut ainsi être pensé "avec" et non "en dépit"du matériau, qui n'est plus neutre et retrouve une fonction harmonique. La rugosité représente alors un moyen de formalisation d'une harmonie non-tonale possédant une réalité perceptive.

Références

[Aures, 1985] Aures, W. (1985). Ein Berechnungsverfahren der Rauhigkeit. *Acustica*, 58:268–281.

[Bigand, 1993] Bigand, E. (1993). The influence of implicit harmony, rythm and musical training on the abstraction of tension-relaxation schemas in tonal music phrases. *Contemporary Music Review*, 9(1 & 2):123–137.

[Bigand et al., 1996] Bigand, E., Parncutt, R., and Lerdhal, F. (1996). Perception of musical tension in short chord sequences : the influence oh harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics*, 58(1):125–141.

[David, 1988] David, H. A. (1988). *The Method of Paired Comparisons*. Oxford University Press, second edition, revised edition.

[Efron, 1981] Efron, B. (1981). Nonparametric estimates of standard error : the jacknife, the bootstrap, and other methods. *Biometrika*, 68:589–599.

[Efron and Tibshirani, 1993] Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall.

[Fletcher, 1940] Fletcher, H. (1940). Auditory patterns. *Rev. Mod. Phys.*, 12:47–65.

[Helmholtz, 1954] Helmholtz, H. L. F. v. (1954). *On the Sensations of Tone as the Physiological Basis for the Theory of Music*. Dover, reprinted from english 2nd ed., 1885 edition. trans. A. J Ellis, from German 4th ed., 1877.

[Hutchinson and Knopoff, 1978] Hutchinson, W. and Knopoff, L. (1978). The acoustic component of western consonance. *Interface*, 7:1–29.

[Kameoka and Kuriyagawa, 1969] Kameoka, A. and Kuriyagawa, M. (1969). Consonance theory part i. *J. Acoust. Soc. Amer.*, 45:1451–1458.

[Lerdahl and Jackendoff, 1983] Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music*. M.I.T Press, Cambridge, MA.

[Moore and Glasberg, 1983] Moore, B. C. J. and Glasberg, B. R. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *J. Acoust. Soc. Amer.*, 74(3):750–753.

[Plomp and Levelt, 1965] Plomp, R. and Levelt, W. (1965). Tonal consonance and critical bandwidth. *J. Acoust. Soc. Amer.*, 38:548–560.

[Schenker, 1935] Schenker, H. (1935). *Der Freie Satz*. Universal Edition, Vienna. Trans. E. Oster (1979) Free Composition, New York: Longman.

[Terhardt, 1974] Terhardt, E. (1974). On the perception of periodic sound fluctuation (roughness). *Acustica*, 30:201–212.

[Zwicker and Fastl, 1990] Zwicker, E. and Fastl, H. (1990). *Psychoacoustics, Facts and Models*. Springer Verlag, Berlin.