

CORRÉLATS ACOUSTIQUES DE CINQ NUANCES DE TIMBRE AU PIANO

Sébastien Bel
Faculté de musique
Université de Montréal
sebastien.bel@umontreal.ca

Caroline Traube
Faculté de musique
Université de Montréal
caroline.traube@umontreal.ca

Résumé

La présente étude a pour objectif de mettre en évidence les corrélats acoustiques de cinq nuances de timbre (brillant, sombre, rond, sec, velouté) produites par quatre pianistes de haut niveau dans le contexte de trois courtes pièces. Autrement dit, nous cherchons à déterminer les descripteurs acoustiques qui permettent de caractériser ces nuances de timbre sur le plan acoustique. Les pièces ayant été jouées sur un piano enregistreur (Bösendorfer CEUS), nous tirons profit de l'enregistrement MIDI du mouvement des touches et du marteau. Nous utilisons ainsi un algorithme permettant d'aligner les données MIDI des prestations enregistrées sur les données MIDI de la partition. Les indices temporels extraits servent à segmenter les captations sonores des prestations. Ensuite, 12 descripteurs acoustiques sont calculés sur chaque segment audio. Finalement, l'analyse en composantes principales est utilisée pour représenter les nuances de timbre dans un espace de paramètres acoustiques permettant de les distinguer. Les résultats mettent en évidence deux dimensions principales, la première étant reliée à l'enveloppe spectrale et la seconde regroupant plusieurs descripteurs d'attaque.

1. CONTEXTE

La question du timbre au piano a longtemps été débattue par la communauté scientifique. Les limitations du contrôle du timbre au piano ont été souvent mise en évidence mais de récents travaux ont montré que les pianistes peuvent percevoir et produire une large palette de nuances timbrales. Pour produire ces timbres, les pianistes ont recours à différents paramètres : la vitesse d'attaque, le contrôle de l'enfoncement de la touche par rapport au seuil de déclenchement du mécanisme de double-échappement, les pédales, l'équilibre dynamique entre les différentes notes jouées simultanément, l'articulation, etc (Goebel et al., 2005). Au sein d'une communauté pianistique

donnée, on peut même retrouver un certain consensus concernant la verbalisation (les termes sont utilisés pour décrire ces timbres) et le geste instrumental à appliquer pour obtenir une nuance de timbre donnée.

La présente étude s'inscrit dans un plus large programme de recherche (initié par le 2e auteur de cet article en 2005) visant à atteindre une compréhension multimodale du timbre au piano, du point de vue du pianiste expert. Elle repose sur des études préalables qui concernent la verbalisation du timbre et l'analyse du geste instrumental et du touché.

Dans le cadre du programme de recherche, des enquêtes menées auprès de pianistes de haut niveau ont d'abord permis de recenser le vocabulaire utilisé pour décrire le timbre au piano (Bellemare and Traube, 2006). Ces enquêtes ont abouti à une liste de près de 100 descripteurs verbaux. Par la suite, des tests de perception ont montré que les interprètes sont capables de reconnaître à l'écoute l'intention timbrale exprimée par un autre pianiste (Bernays and Traube, 2009; Bernays, 2013). Par la technique du regroupement hiérarchique, les nuances de timbre principales ont été déterminées. Ces nuances sont : brillant, sombre, rond, sec, velouté. Un espace de timbre propre au piano a ensuite été obtenu par la technique d'analyse statistique de positionnement multidimensionnel (MDS).

En ce qui concerne l'analyse des paramètres de contrôle appliqués par le pianiste pour obtenir une nuance de timbre donnée, nous avons tiré profit des captations précises du mouvement des touches, du marteau et des pédales que permet le piano acoustique enregistreur Bösendorfer CEUS. De l'analyse de ces données, des *portraits gestuels* des cinq nuances de timbre citées plus haut ont été établis (Bernays, 2013). Simultanément à la captation des paramètres de contrôle au piano, des captations sonores de haute qualité ont été effectuées.

2. OBJECTIFS

La présente étude vise à analyser les captations sonores précédemment effectuées afin de compléter la description multimodale du timbre au piano. Plus précisément, nous avons cherché à mettre en lumière les descripteurs acoustiques qui permettent de caractériser les cinq nuances de timbre étudiées sur le plan acoustique. Ces descriptions complètent ainsi les descriptions verbales et gestuelles obtenues lors des études préalables.

3. MÉTHODOLOGIE

3.1. Enregistrement des nuances de timbre

Pour la collecte des données (paramètres de contrôle et captation sonore), de courtes pièces originales ont été spécialement composées. La consigne qui a été donnée aux compositeurs était de créer des pièces dont la structure ou les mélodies permettraient l'expression de chacune des cinq nuances de timbre étudiées en ne favorisant cependant aucune d'entre elles en particulier. En quelque sorte, les pièces devaient être timbralement neutres. Ces pièces font quelques mesures chacune et les durées ne dépassent pas la vingtaine de secondes, les métriques sont différentes et plusieurs aspects techniques de l'interprétation pianistique y sont inclus.

Ces pièces ont été jouées par quatre pianistes (une femme et trois hommes) de niveau professionnel. Nous les identifierons par leurs initiales dans la suite de l'article (BB, RB, PL, FP). Il est à noter que l'un d'entre eux (FP) possède l'oreille absolue.

Les quatre interprètes ont reçu à l'avance les partitions et les cinq nuances de timbre à exprimer pour pouvoir les répéter. Les répétitions et les enregistrements ont eu lieu au studio du BRAMS (Laboratoire international de recherche sur le Cerveau, la Musique et le Son) sur le piano à queue Bösendorfer Imperial (modèle 290) équipé d'un système d'enregistrement des paramètres de contrôle (mouvement des touches, des pédales et des marteaux). Ces données indexées en fonction du temps sont enregistrées dans des fichiers au format *.boe*.

Lors des sessions d'enregistrement, les pianistes ont reçu la consigne d'interpréter les pièces en leur donnant tour à tour une couleur timbrale parmi les cinq suivantes : brillant, sombre, rond, sec, velouté. La série complète a été jouée trois fois : deux fois dans un ordre de pièces et de timbres choisis par le participant et une fois dans un ordre aléatoire imposé. Le nombre total d'exécutions de ces courtes pièces s'élève ainsi à 240 (4 pièces × 5 nuances de timbre × 4 interprètes ×

3 sessions = 240 enregistrements).

Pièce 1

Frédéric Chiasson



Pièce 2

Stacey Brown



Pièce 4

Frédéric Chiasson



Figure 1: Trois pièces composées pour l'étude sur le timbre au piano

Pour la captation sonore elle-même, deux microphones cardioïdes DPA 4011-TL ont été utilisés pour les prises stéréophoniques proches. Ils ont été disposés en XY à peu près à 1 m du côté ouvert du couvercle du piano, à mi-chemin entre le clavier et le fond de la caisse et à une hauteur de 1,20 m. Les microphones pointaient vers la table d'harmonie du piano.

Deux micros DPA 4006 omnidirectionnels, aussi disposés en XY ont été postés à environ 50 cm derrière les microphones cardioïdes à une hauteur de 1,40 m et pointant au même endroit pour une prise d'ambiance et pour ajouter la réverbération naturelle de la salle dans les enregistrements.

Les signaux des microphones ont été dirigés dans un préamplificateur Millenia HV-2D puis enregistrés

à 44.1 kHz, avec une résolution dynamique de 24 bits. Les quatre pistes ont ensuite été mixées en stéréo avec le logiciel Logic 9[©] et un fichier *wav* à été créé pour chacune des 240 interprétations.

3.2. Segmentation

De façon à pouvoir extraire des descripteurs acoustiques relatifs à l'attaque et à l'évolution temporelle de chaque note ou de chaque groupe de notes jouées simultanément (accords), une segmentation des captations sonores a dû être effectuée dans un premier temps. La segmentation des notes jouées sur un piano est un problème complexe du fait de la nature polyphonique de l'instrument. Plusieurs notes peuvent en effet être jouées simultanément et les notes peuvent également se superposer dans le temps, par l'usage d'une articulation liée ou par l'usage des pédales. La segmentation basée sur l'analyse du contenu spectro-temporel d'une interprétation au piano n'étant pas toujours très robuste, nous exploitons les données de contrôle qui ont été captées par le piano enregistreur (Bösendorfer CEUS) simultanément aux données acoustiques.

Pour effectuer la segmentation de nos captations sonores, nous avons donc dans un premier temps converti les fichiers *.boe* en fichiers MIDI, qui contiennent des informations sur le numéro des touches jouées, sur l'intensité avec laquelle elles sont pressées ainsi que deux informations temporelles relatives au moment initial de l'enfoncement de la touche et à son relâchement (note On / note Off). Dans un deuxième temps, nous avons extrait des fichiers MIDI des partitions encodées dans le logiciel de notation musicale (Sibellius[©]). Ces fichiers MIDI correspondent à une restitution mécanique et régulière de la partition. Finalement, nous avons utilisé un algorithme permettant d'aligner les données MIDI des prestations enregistrées sur les données MIDI de la partition (l'algorithme de *score matching* développé par Large (1993)). De cette manière, chacune des notes de la partition a pu être localisée dans le temps dans chacun des extraits sonores.

L'alignement MIDI n'ayant pas bien fonctionné pour la pièce 3 (vraisemblablement à cause du trille rapide qu'elle contient), il a été décidé d'écarter les données correspondant à cette pièce.

L'ensemble de données analysées contient finalement 180 fichiers audio (pièces 1, 2 et 4) segmentés, événement par événement.

3.3. Calcul des descripteurs acoustiques

11 descripteurs acoustiques ont été sélectionnés parmi les fonctions proposées par la MirToolbox (Lartillot, 2013) dans l'environnement Matlab[©]. Ces descripteurs sont : le centre de gravité spectral (*centroid*), l'étalement spectral (*spread*), l'asymétrie de la distribution spectrale (*skewness*), le *kurtosis* (ou coefficient d'aplatissement de Pearson), le ratio entre la moyenne géométrique et la moyenne arithmétique de la distribution (*flatness*), l'entropie, la *brillance* (reliée à la quantité d'énergie du signal contenue au-dessus de 1500 Hz), le *rolloff* spectral (la fréquence en dessous de laquelle 85% de l'énergie spectrale est contenue), le *zerocross* (le nombre de fois que le signal change de signe), le *temps d'attaque* de chaque événement, la pente moyenne de l'attaque (*attack slope*) et la différence d'amplitude entre le début et la fin de chaque attaque (*attack leap*).

Ces descripteurs sont calculés pour chacun des événements des 180 extraits. Soulignons qu'une interface dédiée a été élaborée (dans Matlab) pour permettre la visualisation des résultats de l'alignement de partition ainsi que de ceux de la segmentation, ou encore pour calculer des descripteurs spécifiques sur des pièces choisies par l'utilisateur. Les retours fournis par cette interface sont des ensembles de données (*datasets*) sur lesquels plusieurs fonctions statistiques de Matlab peuvent être appliquées sans modifications préalables.

Les valeurs moyennes des onze descripteurs acoustiques relatifs à chaque événement sont calculées sur la totalité de la durée des pièces. Chaque pièce est donc représentée par 11 valeurs et les analyses statistiques présentées dans la suite de l'article s'appuient sur ces valeurs moyennes.

4. RÉSULTATS

Les moyennes standardisées des descripteurs pour l'ensemble des 180 interprétations sont représentées sur la figure 2. La représentation en boîte de Tukey permet de comparer la distribution des valeurs de descripteur pour chaque pièce. La marque centrale est la médiane, les bords de la boîte sont les 25e et 75e centiles. On remarque la présence de valeurs extrêmes (représentées par des +) plus nombreuses pour les valeurs qui concernent les descripteurs d'attaques (pentes d'attaques, *attack leaps*, temps d'attaques). Il est possible que ce phénomène soit dû à une segmentation encore trop imprécise pour certaines interprétations.

Une analyse en composantes principales a été appliquée à l'ensemble de données. Comme le montre la

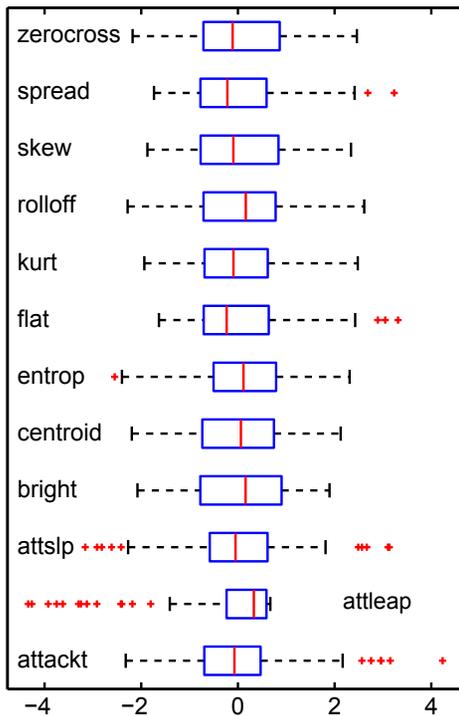


Figure 2: Profil statistique des 11 descripteurs acoustiques pour les 180 interprétations.

figure 3, les deux premières dimensions de la représentation en composantes principales expliquent plus de 70% de la variance. La représentation des données sur les deux premières dimensions, soit sur deux axes, devrait donc être satisfaisante.

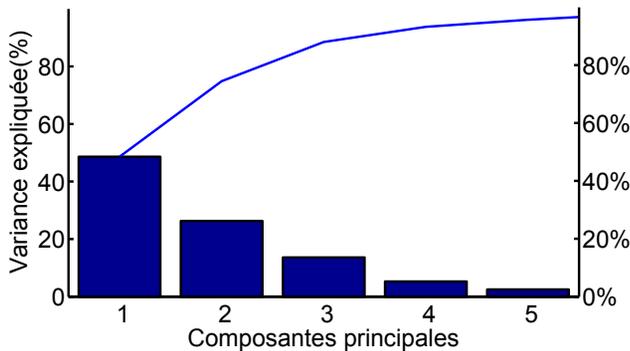


Figure 3: Représentation de la variance expliquée en fonction de la composante principale.

Par ailleurs, on voit sur la figure 4 que ces deux dimensions sont de natures différentes, puisque la première oppose essentiellement des descripteurs spectraux (*zerocross*, *bright*, *rolloff*, *centroid* opposés à *kurtosis* et *skewness*) et qu'une seconde dimension semble

opposer des descripteurs d'attaque (*attcktime* opposé à *attckslp* et *attleap*) ainsi que certains descripteurs spectraux (*flatness*, *spread*, *entropy*). Il est aussi intéressant de noter que la première dimension dépend fortement des descripteurs *centroid*, *rolloff* et *brightness*, ce qui semble correspondre avec les résultats de nombreuses études démontrant l'importance du centre de gravité spectral et d'autres corrélats acoustiques de la brillance, dans la perception des différences de timbres instrumentaux.

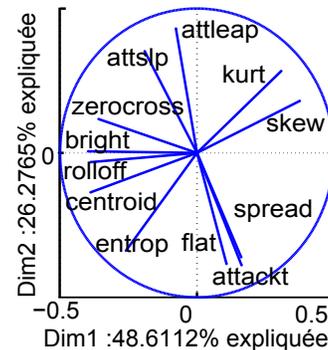


Figure 4: Contribution des différents descripteurs à la représentation en composantes principales.

Dans la suite de cet article, les dimensions 1 et 2 seront représentées sur l'axe horizontal et l'axe vertical respectivement. Ainsi, l'axe horizontal est relié aux caractéristiques spectrales des événements sonores alors que l'axe vertical est relié aux descripteurs d'attaque des événements.

Sur l'axe vertical on constate que les composantes *penste de l'attaque* et *hauteur de l'attaque* sont opposées à la composante *temps d'attaque*, ce qui semble cohérent dans la mesure où cela oppose les sons d'attaques longues avec les sons dont la pente d'attaque est forte (donc dont l'attaque est rapide).

À présent, si on regarde la représentation en composantes principales des descripteurs acoustiques des 180 interprétations (Fig. 5), on constate que les moyennes générales de chaque nuance (représentées par les gros polygones) sont réparties avec un étalement assez grand sur l'axe horizontal, ce qui semble montrer qu'elles se distinguent assez bien spectralement sauf pour les nuances brillant et sec qui sont assez proches pour les trois pièces. En revanche, l'axe vertical qui représente les descripteurs d'attaque est beaucoup moins variable (les points moyens sont beaucoup plus proches les uns des autres sur l'axe vertical).

La faible variabilité sur l'axe vertical (axe des descripteurs d'attaque) aurait tendance à montrer que les différentes nuances de timbre ne sont pas dues

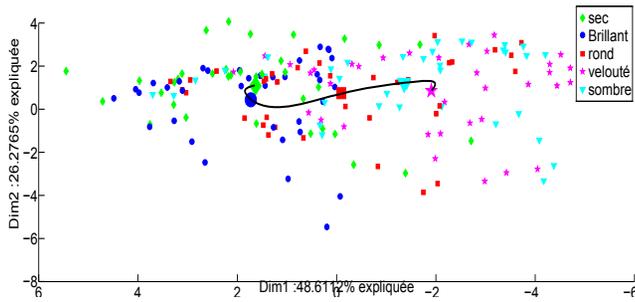


Figure 5: Analyse en composantes principales de 11 descripteurs acoustiques sur les 180 interprétations que composent les pièces 1,2 et 4.

principalement à des différences dans l'attaque des sons, mais à des différences spectrales. En revanche, les nuances brillant et sec sont acoustiquement très proches (sur l'axe horizontal), seul l'axe vertical (temporel) semble les distinguer –on peut par exemple observer ce phénomène dans la représentation des interprétations de la première pièce (voir Fig. 6).

Si on tient compte des significations respectives des deux dimensions évoquées plus haut (fig.4), et des positions respectives des interprétations moyennes de chaque nuance (représentées par des points plus gros) on peut comprendre que le timbre sec correspond à un son dont l'attaque est plus forte en intensité (*attack slope* et *attack leap* ont plus d'importance) et moins longues (*attack-time* à moins d'importance). Cette disposition montre que les pianistes, lorsqu'on leur demande de produire un timbre sec, produisent par leurs gestes des sonorités proches spectralement des sons brillants, mais dont les attaques sont plus brèves et plus intenses.

Puisque chaque pianiste a eu trois essais pour interpréter chaque nuance de timbre, nous avons calculé un point représentant une interprétation moyenne des trois essais de chaque interprète pour chaque nuance. Par l'observation des graphiques à la figure 6, on peut constater que les interprétations moyennes d'un pianiste varient fortement le long de la première dimension et très peu sur le long de la seconde dimension. Cela est mis en évidence sur le troisième graphique correspondant à la pièce 4, où les interprétations moyennes sont reliées entre elles par une courbe, pour chacun des quatre interprètes. Ces courbes se distribuent du bas vers le haut dans l'ordre suivant : PL, FP, BB, RB. On peut également remarquer que l'étalement de cette courbe varie de façon importante d'un interprète à l'autre (BB<FP<PL<RB).

Les pianistes semblent donc s'accorder sur la disposition relative des nuances du point de vue spectral (de gauche à droite : sec, brillant, rond, sombre, ve-

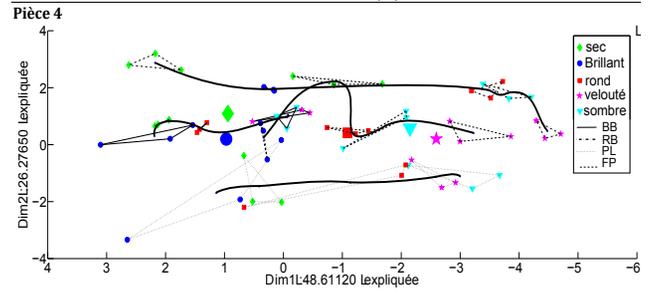
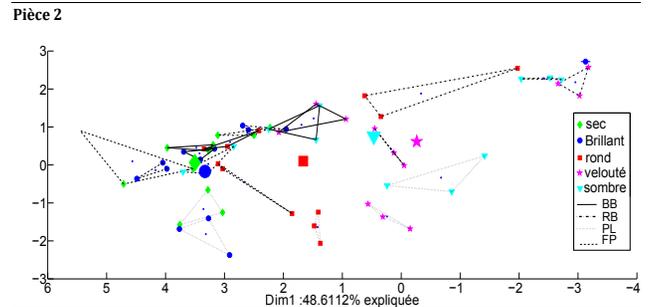
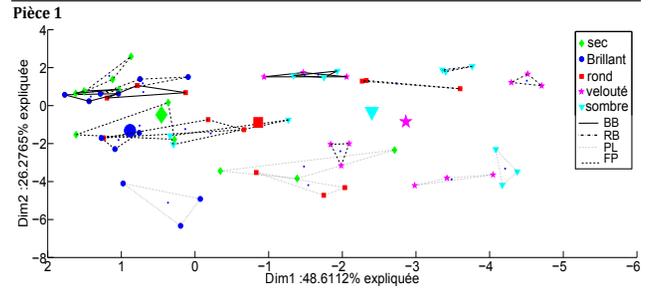


Figure 6: Analyse en composantes principales de 11 descripteurs acoustiques sur les 60 interprétations de la pièce 1, 2 et 4

louté), alors que leur individualité s'exprime davantage par un positionnement distinctif le long de l'axe des descripteurs d'attaques (axe vertical) et par l'étendue de la palette spectrale qu'ils sont capables de produire (représentée par la longueur de la courbe).

De façon plus générale, on remarque que trois nuances sont disposées dans le même ordre sur l'axe horizontal pour les trois pièces : velouté, sombre, rond. Sur l'axe vertical, ces couleurs se disposent comme suit : rond est toujours au-dessous de sombre, qui est lui-même toujours au-dessus de velouté (voir Fig. 5, 6).

Le fait que les positions relatives de ces nuances d'un point de vue spectral (axe horizontal) et d'un point de vue temporel (axe vertical) sont conservées peu importe la pièce suggère qu'il existe dans l'espace acoustique des sous-espaces de valeurs de descripteurs spectraux et temporels propres à certaines colorations timbrales.

Pour les deux dernières nuances, brillant et sec, cette conservation des positions relatives est moins claire sur l'axe spectral, car les nuances sont très proches le long de l'axe horizontal dans les trois pièces. En revanche, le long de l'axe vertical, il semble y avoir conservation de leur position relative. Cela concorde avec ce que nous entendons des pièces jouées avec la nuance de timbre «sec», leurs événements présentant des attaques plus courtes et d'intensités légèrement plus grandes que les interprétations avec le timbre «brillant».

5. DISCUSSION ET CONCLUSION

Les résultats présentés ici fournissent une première description acoustique des cinq colorations timbrales (brillant, sombre, velouté, rond et sec) étudiées au préalable du point de vue du contrôle gestuel au niveau du clavier et des pédales (Bernays, 2013). Les nuances de timbre semblent se distinguer et se répartir le long d'un premier axe dont les composantes les plus influentes sont des corrélats acoustiques de la brillance. Par ailleurs, l'individualité des pianistes s'exprime par un positionnement le long de l'axe vertical qui relié à des descripteurs d'attaque.

Ces résultats convergent avec ceux de l'étude sur le geste producteur de timbre au piano mentionnée plus haut. Dans cette étude, la même banque de 180 interprétations est analysée en fonction de descripteurs gestuels et articulatoires (vitesse d'attaque, degré de synchronisme, degré de legato, etc.). Une analyse en composantes principales aboutit à une représentation graphique de l'espace de timbres suivant des dimensions reliées au geste instrumental et à l'articulation qui est semblable à celle que nous obtenons ici dans l'espace des descripteurs acoustiques. Cette disposition est aussi constatée lors de l'arrangement des cinq descripteurs de timbres dans l'espace sémantique – à l'exception des nuances sombre et velouté dont les positions sont inversées (Bernays, 2013).

Il est possible que la faible répartition des nuances sur l'axe vertical (dont la plupart des composantes sont reliées à l'attaque ou à l'intensité) provienne d'une part du fait que la vitesse d'attaque et de la dynamique d'une note au piano sont fortement interdépendantes et d'autre part, du fait que les interprètes aient reçu pour consigne de limiter les variations de dynamique (le timbre étant par définition différent de l'intensité). Ainsi, il semble assez normal que les descripteurs *attack leap*, *pente de l'attaque* et *durée d'attaque* prennent des valeurs assez proches pour un interprète donné.

Notons que certains problèmes de segmentation des fichiers audio nous ont empêché d'appliquer la

méthodologie proposée à la pièce 3 et causent encore certaines erreurs dans les calculs de descripteurs. Ces erreurs proviennent de la difficulté à appliquer la segmentation sur des ensembles de notes courtes et proches dans le temps (typiquement des trilles). Pour la poursuite de cette étude, nous envisageons d'ajouter des descripteurs acoustiques issus d'autres boîtes à outils développées sous Matlab[®] qui sont complémentaires à ceux ici exploités (les descripteurs acoustiques reliés à la chute du son par exemple).

En conclusion, par cette étude et celles qui la précèdent, le phénomène du rapport geste-son au piano est cerné suivant plusieurs perceptives complémentaires : celle du geste (quel geste est appliqué sur l'instrument), celle du son (quels caractéristiques acoustiques possèdent le son produit), et celle de la perception (comment ces variations sonores sont-elles perçues). Par la présente étude, nous espérons contribuer à une meilleure compréhension du rôle joué par plusieurs caractéristiques physiques du signal acoustique dans la perception qu'ont les pianistes des nuances de timbre qu'ils produisent en contexte musical.

Références

- BELLEMARE, Madeleine and TRAUBE, Caroline (2006). Investigating piano timbre : Relating verbal description and vocal imitation to gesture, register, dynamics and articulation. In *International Conference on Music Perception and Cognition*. Bologne, 59–60.
- BERNAYS, Michel (2013). *Expression and production of piano timbre : Gestural control and technique, perception and verbalisation in the context of piano performance and pedagogy*. PhD thesis.
- BERNAYS, Michel and TRAUBE, Caroline (2009). Expression of piano timbre : gestural control, perception and verbalisation.
- GOEBL, Werner, BRESIN, Roberto, and GALEMBO, Alexander (2005). Touch and temporal behavior of grand piano actions. *The Journal of the Acoustical Society of America*, 118(2) :1154–65. URL <http://www.ncbi.nlm.nih.gov/pubmed/16158669><http://link.aip.org/link/JASMAN/v118/i2/p1154/sl&Agg=doi>.
- LARGE, Edward W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, and Computers*, 25(2) :238–241. URL <http://link.springer.com/article/10.3758/BF03204504>.
- LARTILLOT, Olivier (2013). MIRtoolbox 1.4 .1.