



HAL
open science

L'apport des technologies de la musique pour la recherche en neurosciences

Jean-Julien Aucouturier

► **To cite this version:**

Jean-Julien Aucouturier. L'apport des technologies de la musique pour la recherche en neurosciences. 2016. hal-01450515

HAL Id: hal-01450515

<https://hal.science/hal-01450515>

Submitted on 20 Nov 2019

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.

L'apport des technologies de la musique pour la recherche en neurosciences

JJ Aucouturier

La musique détient un bien étrange pouvoir sur nos émotions. Au détour d'une phrase, d'une texture, survient l'événement sonore, une attaque, un tremblement, un seul soupir parfois, qui met toute notre physiologie en aboi. Pour le neuroscientifique, ces événements, ces signaux sont de vexantes provocations, car si l'on crie, si notre cœur s'emballe et si nos mains deviennent moites en tombant nez à nez avec un ours lors d'une balade en forêt, il sait bien que c'est pour des raisons de survie: le cri avertit nos semblables du danger, le cœur pompe plus de sang dans nos muscles pour nous préparer au combat, et notre transpiration régule notre température pour nous préparer à courir. Tomber nez à nez avec de la musique, avec un do ou un do#, un violon ou une flûte, est rarement une question de vie ou de mort, et pourtant nos réactions biologiques sont les mêmes.

Une des hypothèses les plus contemporaines sur cette question est que notre cerveau se ferait, en fait, "avoir" par la musique, qui se fait passer pour un stimulus plus important qu'elle n'est en réalité. Car il est bel-et-bien des sons auxquels il est important de répondre vite et intensément: le rugissement dissonant d'un fauve, l'éclat de voix colérique d'un rival, le rire contagieux d'un enfant ou la berceuse apaisante d'une mère. Avec ses phrases hésitantes ou nettement scandées, ses timbres sombres ou brillants, ses notes tremblantes ou soutenues, il se peut donc que la musique active, à notre insu, les circuits neuronaux normalement recrutés pour traiter la parole expressive, et déclenche ainsi des réactions émotionnelles usurpées. Malheureusement, pour tester cette idée, les outils manquent: il faudrait pouvoir produire de la musique qui "sonne" comme de la voix émotionnelle, dont les caractéristiques acoustiques reproduiraient précisément la délicate harmonie d'intensité, hauteur, timbre et durée de la voix humaine quand elle se fait joyeuse, triste ou inquiète. Faute de tels outils, les neurosciences de la musique en sont réduites à constater que la musique et la voix activent parfois les mêmes zones, mais parfois non, parfois les mêmes émotions, mais parfois non, sans véritablement comprendre pourquoi ni comment.

Si cette question de faire parler un orchestre, ou de faire chanter une parole, reste en somme pour les neurosciences une hypothèse en quête d'outil, il est un autre domaine où, paradoxalement, les exemples pratiques de telles réalisations abondent: celui de l'informatique musicale et de la création musicale contemporaine. Dans les studios de musique de l'IRCAM, depuis 40 ans, des compositeurs comme Luciano Berio, Jonathan Harvey ou Salvatore Sciarrino resoumettent régulièrement la question aux ingénieurs, chercheurs et réalisateurs en informatique musicale de comment utiliser l'ordinateur pour transcrire en sons musicaux l'essence d'une expression vocale (par ex. Paul Scofield lisant TS Eliot imité par un trombone à coulisse dans *Speakings* de Jonathan Harvey). Les résultats de ces tentatives, parfois rapidement retournés dans les cartons ou au contraire développés au point d'être réutilisés depuis dans d'autres compositions, ne sont souvent jamais sortis des studios, et n'ont jamais été confrontés à la question scientifique qu'ils permettraient pourtant de résoudre. Pour les neurosciences de la musique, l'IRCAM est une vraie caverne d'Ali Baba, une mine d'or qui ne demande qu'à être exploitée.

Grâce à un généreux soutien de l'European Research Council, j'ai fondé l'équipe CREAM (« Cracking the Emotional Code of Music ») dans le laboratoire STMS de l'IRCAM précisément dans ce but : développer l'apport des technologies du traitement du signal sonore pour l'expérimentation en psychologie et neurosciences cognitives. Notre ambition est d'avancer sur ces deux fronts (développer des outils logiciels pour la communauté neuro et utiliser ces logiciels pour nos propres travaux de neuroscience) et de positionner le laboratoire comme pionnier d'une méthodologie qui, je l'espère, changera radicalement la façon dont on étudie comment le cerveau traite la musique et le son.

Premier exemple de cette démarche : nous venons de publier cette année un logiciel open-source, DAVID, capable de transformer le ton émotionnel de la voix, en temps-réel. S'inspirant d'une longue tradition ircamienne d'analyse/synthèse de la voix, DAVID est capable en moins de 20ms de s'emparer d'un flux de paroles prononcées au micro, et d'en modifier algorithmiquement la prosodie et le timbre pour les faire paraître plus joyeuses, tristes ou anxieuses qu'elles ne le sont en vérité. Nous avons validé l'effet avec des expériences d'écoute dans 4 pays (France, Angleterre, Suède et Japon) qui ont montré que les transformations sont tellement réalistes que les locuteurs eux-mêmes pensent qu'il s'agit de leur propre voix. Comme l'effet fonctionne en temps-réel, les applications scientifiques sont nombreuses : il devient par exemple possible d'étudier en laboratoire des situations de dialogue entre deux personnes dont on modifie à loisir le ton de la voix - est-on plus créatifs en groupe quand tout le monde parle avec la même émotion ? Préfère-t-on une solution quand on s'entend l'énoncer avec une voix joyeuse ?

Dans une première étude utilisant DAVID, nous avons demandé à des participants de lire un texte à voix haute en s'écoutant au casque. A leur insu, nous avons modifié leur voix pour la faire paraître plus joyeuse ou plus triste. Non seulement les participants n'ont pas détecté la manipulation, mais leur état émotionnel a été modifié dans la direction de la transformation : ceux qui s'entendaient parler plus gaiement sont devenus plus gais, et les autres, plus tristes. Ce résultat, que nous venons de publier dans les compte-rendus de l'Académie Américaine des Sciences (PNAS), est important sur le plan fondamental pour les neurosciences de l'émotion car il constitue la première démonstration d'une rétroaction émotionnelle dans le domaine de la voix. D'un point de vue applicatif, il ouvre des perspectives cliniques dans le traitement de la dépression et des troubles post-traumatiques (voir entretien ci-contre avec Philippe Fossati). De façon plus générale, il illustre parfaitement la raison d'être de l'équipe CREAM : voilà une découverte neuroscientifique qui n'aurait jamais pu voir le jour sans l'expertise en traitement du signal vocal et musical de l'IRCAM. La première d'une longue série, on espère...

Liens :

Equipe CREAM : <http://cream.ircam.fr>

Le Logiciel DAVID : [lien à venir]

Etude PNAS : <http://www.pnas.org/content/113/4/948.full.pdf> en collaboration avec le LEAD de l'Université de Bourgogne à Dijon, le Département de Sciences Cognitives de l'Université de Lund en Suède et les Universités de Tokyo et de Waseda au Japon.