Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique

Université de Sfax

institut supérieur de musique de sfax

LARIDIAME
Laboratoire de Recherche Interdisciplinaire en Discours, Art, Musique et Economie

المؤتمر الدولي الثامن حول تحليل الخطاب الموسيقي

توظيف الذكاء الإصطناعي في تأليف الموسيقات المقامية

COLLOQUE INTERNATIONAL N°8 :

CRÉATION MUSICALE ET INTELLIGENCE ARTIFICIELLE EN MUSIQUES DU MAQĀM

إشراف وتنسيق: أ.د. الأسعد الزواري

د. حلمي بنصير

CRT Sfax
Le centre de ressources technologiques

09-10 MARS 2020

laridiame.mus@gmail.com

WWW.ISMSF.RNU.TN

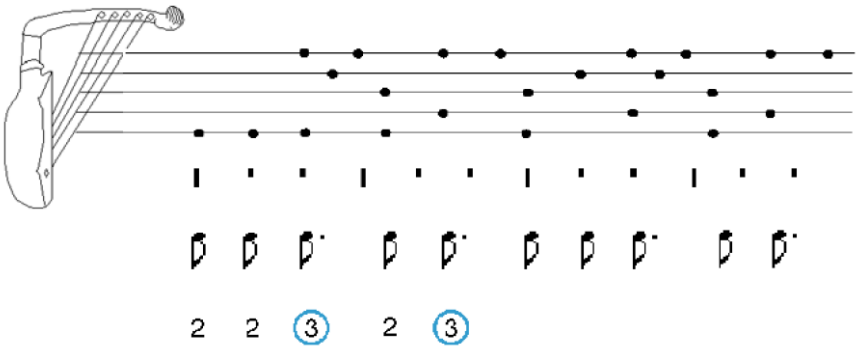
De la musique aux mathématiques... et réciproquement

Marc Chemillier
marc.chemillier@ehess.fr

Au cours de cet exposé, je vais parler de mathématiques et de musique de deux points de vue différents qui sont, pourrait-on dire, complémentaires. D'une part, je vais montrer comment la musique fait apparaître certaines structures remarquables que l'on peut analyser en utilisant des outils mathématiques. Et dans un deuxième temps je montrerai comment les mathématiques peuvent proposer des modèles qui permettent de générer de la musique et de simuler des musiques existantes à l'aide de techniques de l'intelligence artificielle. On verra que dans ce cas, l'approche générative conduit à des applications intéressantes pour enrichir l'analyse musicale.

Le premier exemple que je vais traiter apparaît dans le contexte un peu inattendu d'une société de tradition orale. J'ai en effet travaillé en Afrique où l'on trouve certaines structures musicales tout à fait étonnantes que l'on peut décrire mathématiquement. C'est le cas d'un répertoire d'Afrique centrale sur lequel j'ai travaillé dans les années 1990, chez les Nzakara qui jouent de la harpe pour accompagner le chant. Le premier exemple reproduit ci-dessous est une petite formule de harpe qui est jouée en boucle de façon répétitive avec quelques variations, et qui sert d'accompagnement à une improvisation vocale (il s'agit de la pièce *gitangi* dans le CD de la Cité de la musique). Sur le plan du rythme, l'auditeur est un peu déstabilisé parce qu'il a l'impression que le jeu de harpe flotte en quelque sorte alors qu'en réalité, il est très structuré du point de vue rythmique. Pour savoir comment cette formule instrumentale est organisée, il faut déterminer les positions de la pulsation. La méthode pour le faire consiste à demander à des gens du village de frapper dans les mains en écoutant la séquence. C'est ce que

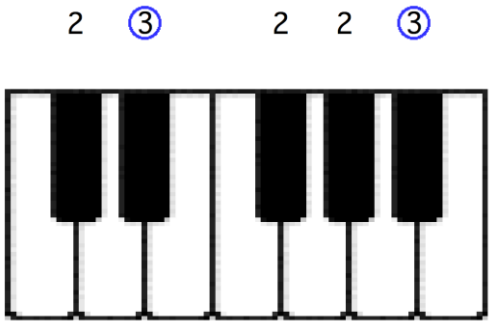
j'ai fait pour obtenir le résultat présenté dans la figure ci-dessous. Les pulsations se placent aux positions des traits verticaux les plus longs. Il y en a quatre dans la séquence avec trois subdivisions pour chaque pulsation. Quand on regarde les « notes » de musique, c'est-à-dire les points où le musicien pince les cordes de l'instrument, on voit qu'au début ils coïncident avec la pulsation ou avec ses subdivisions et qu'ensuite, ils se placent entre les traits verticaux flottant en l'air quelque sorte. C'est ce qui donne le caractère rythmique très particulier de la séquence. On peut décrire les rapports de durées de manière plus précise avec la notation solfégique. Il suffit de représenter la subdivision de la pulsation par une croche qui se divise en deux double-croches. Par endroits la durée de certaines notes est un peu prolongée, c'est-à-dire que la note suivante arrive plus loin, et cela correspond à une croche pointée, c'est-à-dire une croche plus la moitié de la durée d'une croche.



En observant la suite de durées qui en résulte on obtient une séquence très particulière que l'on peut représenter mathématiquement en notant les croches par le chiffre 2 (car il y a deux subdivisions en double-croches) et les croches pointées par le chiffre 3. Cela donne la succession 2 2 3 2 3 qui est répétée ensuite. C'est cette structure qui donne à ce rythme son caractère si particulier avec par moments des attaques qui coïncident avec la pulsation ou avec ses subdivisions, et à d'autres des attaques qui sont en levée, c'est-à-dire entre les subdivisions de la

pulsation, par exemple entre la deuxième et la troisième pulsations ou entre la quatrième et la fin de la séquence. Cette propriété est une caractéristique des rythmes africains. Il existe plusieurs rythmes de la même famille qui ont exactement la même organisation des durées avec des 3 séparées par des groupes de 2 comportant un élément de plus d'un côté que de l'autre, donc une dissymétrie dans le formule. C'est ce qui donne le caractère « syncopé » de ces rythmes, ce côté très rebondissant propre aux musiques africaines. Ce sont de telles structures que l'on peut étudier mathématiquement et le recours aux mathématiques est d'autant plus intéressant qu'on trouve des structures similaires dans des contextes complètement différents comme on va le voir sur un exemple.

Si l'on reprend la succession 2 3 2 2 3 avec toujours l'asymétrie dans les groupes de 2, il y a un objet très familier en musique dans lequel on retrouve exactement la même disposition, il s'agit du clavier du piano. La répartition des touches noires par rapport aux touches blanches suit exactement la même structure. Entre le *do*♯ et le *ré*♯, on a deux touches (une noire et une blanche). Par contre entre le *ré*♯ et le *fa*♯, on a trois touches (*ré*♯, *mi, fa*). Ensuite, on trouve de nouveau deux touches, puis deux touches, puis trois touches entre le *si*♭ et le *do*♯. Donc on a exactement la même organisation dans la répartition des touches du clavier.



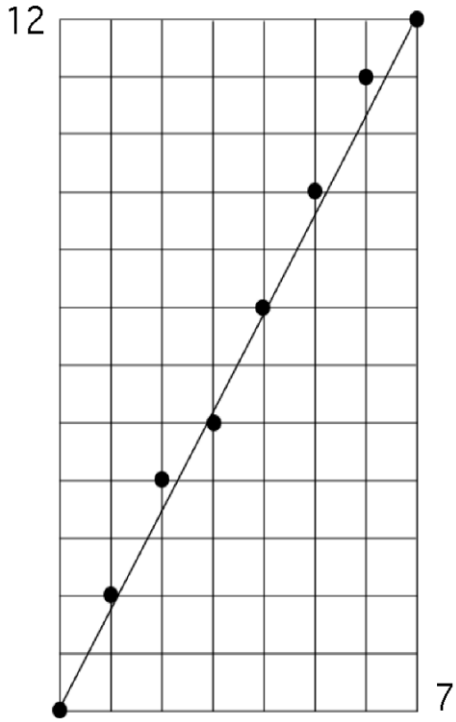
Ce type de structure asymétrique a donné lieu à des explorations mathématiques. Sur le plan des rythmes, c'est le mathématicien Godfried Toussaint qui a publié en 2013 un livre sur les patterns asymétriques que l'on trouve en Afrique mais aussi dans d'autres contextes comme les Caraïbes (claves afro-cubaines) ou l'Europe centrale (rythmes aksak).

En ce qui concerne les touches du clavier, c'est-à-dire les gammes et les modes, la musicologie américaine s'est beaucoup intéressée à ce sujet en développant le concept de *sous-ensemble maximalement réparti* par rapport à un ensemble donné dont il existe plusieurs caractérisations mathématiques. Nous allons en donner une qui est assez facile à exprimer. On représente un quadrillage avec sept unités en largeur et douze unités en hauteur et on trace la droite diagonale de pente 12/7. Précisons que les nombres 7 et 12 s'expliquent par le fait qu'il y a sept touches blanches à l'intérieur de l'octave et douze touches dans le total chromatique (sept blanches et cinq noires). On cherche les points d'intersection du quadrillage « les plus proches » de la droite, dans un sens que l'on va préciser ultérieurement, et l'on s'aperçoit qu'on trouve exactement la succession des touches blanches du clavier : *do, ré, mi*, puis entre le *mi* et le *fa* on voit qu'il y a seulement une unité (alors qu'il y en avait deux entre les notes précédentes). À partir du *fa* on trouve de nouveau deux unités entre les notes *fa, sol, la, si*, puis à la fin une seule unité entre le *si* et le *do*. La formule qui permet de caractériser ces points de coordonnées (i, j) est :

$$j = E((i + 1)p) - 1$$

dans laquelle E est la fonction partie entière qui donne l'entier le plus proche inférieur à un nombre donné, et $p = 12/7$ est la pente de la droite. Cette formule linéaire donne exactement les sept touches blanches du clavier à l'intérieur des douze divisions de l'octave. On voit bien que si l'on avait eu un 6 à la place du 7, les choses auraient été très différentes parce que 12 est divisible par 6 et dans ce cas, la droite de pente 12/6 = 2 aurait croisé le quadrillage en des points de coordonnées entières. On aurait

obtenu une belle régularité avec des points espacés de deux unités (cela correspond à la gamme par ton sur le clavier). Mais la nature en a voulu autrement puisqu'il y a sept touches blanches. On a donc cette structure représentée sur la figure qui est « presque » régulière, qui est la plus régulière possible compte tenu du fait que 12 et 7 sont des nombres premiers entre eux. C'est cela qui caractérise ces ensembles maximalement répartis : ils sont les plus proches possible d'ensembles répartis de façon régulière.



Lorsqu'on veut énumérer de telles structures, qu'elles soient rythmiques ou modales, on se trouve confronté au problème de la circularité de ces séquences. Par exemple si l'on prend la gamme pentatonique correspondant aux touches noires du clavier, elle est définie par la suite d'intervalles 2 3 2 2 3 lorsqu'on commence par le do♯. Mais toutes les rotations de cette séquence sont également des gammes pentatoniques. On obtient

en effet cinq formes différentes de la gamme pentatonique en commençant par chacune de ses notes *do*♯, *ré*♯, *fa*♯, *sol*♯, *si*♭ :

2 3 2 2 3
3 2 2 3 2
2 2 3 2 3
2 3 2 3 2
3 2 3 2 2

Il peut exister des circonstances spécifiques dans lesquelles ces différentes rotations doivent être distinguées. Mais lorsque l'on parle de la gamme pentatonique en général, on ne les distingue pas et elles sont toutes considérées comme équivalentes. Dans ce cas, il est utile de disposer d'une technique pour sélectionner un représentant unique parmi les différentes rotations. Un moyen simple de le faire est de considérer la rotation qui est *la plus petite pour l'ordre alphabétique*. Ainsi pour la gamme pentatonique, on obtient la rotation 2 2 3 2 3 (on place les chiffres les plus petits en positions les plus à gauche).

Ce représentant unique qui est minimal pour l'ordre alphabétique parmi les rotations d'une même séquence est appelé un *mot de Lyndon*. C'est un concept très puissant de combinatoire des mots et il existe un algorithme permettant de les calculer très efficacement (algorithme de Duval). Dans son étude sur la systématique modale, Amine Beyom distingue la notation « prescriptive » d'un mode qui sert à définir la structure de l'échelle et permet de dessiner une carte du territoire des musiques modales, et la notation « descriptive » qui cherche à rendre compte des variations d'intonation servant à différencier les pratiques géographiques, personnelles, etc. Il donne l'exemple du mode Bayāt qui utilise le demi-bémol noté ♭ : *ré, mi*♭, *fa, sol, la, si*♭, *do, ré*. Le degré *mi*♭ forme un intervalle de 3 quarts de ton avec le *ré* qui précède et avec le *fa* qui suit. Dans la pratique, c'est un degré appelé SĪKĀ qui peut varier selon les musiciens et les régions. Mais du point de vue de la structure de l'échelle, la notation prescriptive retiendra la séquence d'inter-

valles suivante entre les notes de l'échelle considérée en partant de la note *ré* :

3 3 4 4 2 4 4

La systématique modale d'Amine Beyom a permis une énumération rationnelle des structures d'intervalles de la modalité arabe. Celles-ci sont représentées comme des séquences de 7 intervalles multiples du quart de ton. Il distingue l'*hyper-système* qui est un ensemble de 7 intervalles placés en ordre croissant : pour le Bayāt, il s'agit de 2 3 3 4 4 4 4. Les permutations que l'on peut former avec les hyper-systèmes sont appelées des *sous-systèmes*. Parmi eux, on choisit un représentant unique pour chaque famille de sous-systèmes identiques à rotation près. Ce représentant unique, appelé *système,* est la séquence minimale pour l'ordre alphabétique parmi toutes les rotations. C'est donc un mot de Lyndon et l'algorithme de Duval permet de les calculer. On obtient le calcul suivant pour les séquences de longueur 7 qui sont basées sur l'hyper-système 2 3 3 4 4 4 4 du mode Bayāt :

```
(duval-hyper 7 '(2 3 4) #(2 3 3 4 4 4 4))
lyndon=#(2 3 3 4 4 4 4)
lyndon=#(2 3 4 3 4 4 4)
lyndon=#(2 3 4 4 3 4 4)
lyndon=#(2 3 4 4 4 3 4)
lyndon=#(2 3 4 4 4 4 3)
lyndon=#(2 4 3 3 4 4 4)
lyndon=#(2 4 3 4 3 4 4)
lyndon=#(2 4 3 4 4 3 4)
lyndon=#(2 4 3 4 4 4 3)
lyndon=#(2 4 4 3 3 4 4)
lyndon=#(2 4 4 3 4 3 4)
lyndon=#(2 4 4 3 4 4 3)
lyndon=#(2 4 4 4 3 3 4)
lyndon=#(2 4 4 4 3 4 3)
lyndon=#(2 4 4 4 4 3 3)
```

On trouve 15 systèmes associés à l'hyper-système 2 3 3 4 4 4 4 ce qui confirme le tableau d'énumération publié par Amine Beyom (p. 10 de son article). Le mot de Lyndon correspondant au mode Bayāt est le dixième de la liste : 2 4 4 3 3 4 4 qui représente le mode considéré à partir de la note *la*, c'est-à-dire *la, si♭, do, ré, mi♭, fa, sol, la*. Les autres permutations correspondent à des modes différents.

Lorsque des objets combinatoires remarquables apparaissent dans des sociétés de tradition orale, comme c'est le cas pour les rythmes africains, on peut se demander quels sont les processus cognitifs mis en œuvre pour permettre l'émergence de telles structures. C'est la question que je me suis posée dans mon livre *Les Mathématiques naturelles* qui est une exploration d'une discipline nouvelle qu'on appelle l'ethnomathématique et qui consiste à étudier ce genre de structure, non seulement dans la musique mais également dans de nombreuses activités pratiquées dans les sociétés de tradition orale.

Je vais passer maintenant à l'autre sujet de mon exposé qui est complémentaire en quelque sorte, puisqu'il s'agit de voir comment les mathématiques permettent de générer de la musique avec certains modèles comme les automates finis ou les chaînes de Markov. Je vais prendre de nouveau un exemple pour expliquer le principe. Il s'agit d'étudier des séquences qui peuvent contenir des notes si l'on parle de musique, mais qui dans le cas général peuvent être des séquences d'objets quelconques. Pour illustrer cette construction, prenons l'exemple d'une séquence de lettres : *abracadabrabradaca*. On va se demander dans un contexte donné, c'est-à-dire quand on a lu un certain nombre de lettres, qu'est-ce qu'on peut lire à la suite en se limitant aux enchaînements définis par la séquence de départ ? Dans la séquence qui est proposée, si l'on prend comme contexte les quatre lettres *abra*, on constate qu'il y a trois transitions possibles : soit on peut lire un *c*, soit un *b*, soit un *d*. En effet, il y a trois occurrences du contexte *abra* dans la séquence avec à chaque fois une suite différente. Voilà ces trois occurrences indiquées entre parenthèses :

(*abrac*)*adabrabradaca*
abracad(*abrab*)*radaca*
abracadabr(*abrad*)*aca*

Le point intéressant à observer est que si l'on modifie la longueur du contexte, par exemple en l'augmentant par ajout d'une lettre (en prenant *rabra* à la place de *abra*), on réduit les possibilités de choix de transitions. Alors qu'au contraire, si on diminue la longueur du contexte, on augmente le nombre de choix possibles. Par exemple si l'on prend *rabra*, on s'aperçoit que dans cette séquence, il n'y a plus qu'une seule transition possible vers le *d*, donc on a réduit l'éventail des possibilités :

abracadab(*rabrad*)*aca*

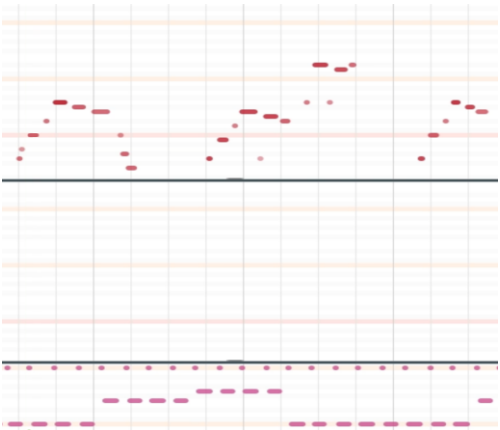
On peut utiliser ce modèle dans un sens génératif, c'est-à-dire en produisant des séquences par un tirage effectué au hasard selon les probabilités de transition de la séquence enregistrée initialement pour produire une séquence qui lui ressemble beaucoup. Voilà un exemple de séquence obtenue en faisant varier le contexte : *dacadabrabracada*. On a indiqué ci-dessous les étapes qui permettent de produire cette séquence en utilisant des parenthèses pour délimiter les contextes les plus longs appartenant à la séquence d'origine *abracadabrabradaca* :

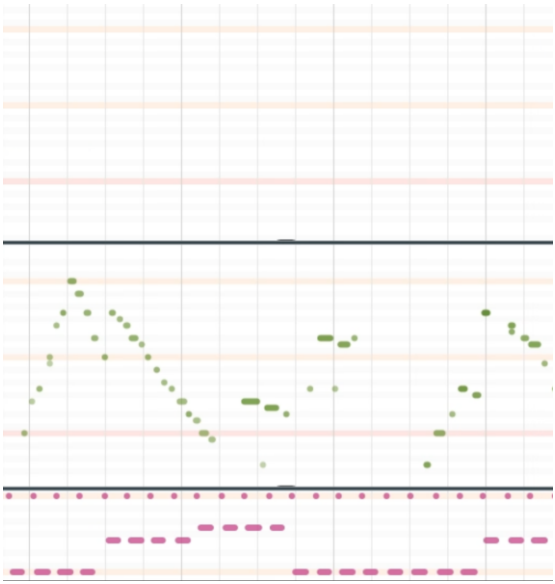
(*daca*)
d(*acadabrabra*)
dacadabr(*abracada*)

On voit que le contexte augmente lorsqu'on suit l'ordre des lettres dans la séquence d'origine, puis il se rétrécit lorsqu'on bifurque et dans ce cas, on « s'éloigne » en quelque sorte de la séquence initiale. Ce rétrécissement est visible lorsqu'on passe de *d*(*acadabrabra*) à *dacadabr*(*abrac).* Puis lorsqu'on augmente de nouveau le contexte, on se rapproche de la séquence initiale et ainsi de suite. Quand le contexte augmente, on a tendance à recopier des fragments de ce qui a été enregistré, et c'est quand

le contexte diminue que l'on crée de la nouveauté.

On utilise ces principes dans un travail que j'ai initié avec des collègues de l'IRCAM il y a une vingtaine d'années pour le développement d'un logiciel d'improvisation qui est à un stade assez avancé puisqu'on l'utilise désormais en concert dans différents projets artistiques avec le musicien de jazz Bernard Lubat ou le musicien Charles Kely Zana-Rotsy de Madagascar. Je vais montrer pour finir un exemple avec Bernard Lubat illustrant le fonctionnement du logiciel qui est basé exactement sur ce principe (https://youtu.be/J52SEoHvx1o). On capte ce que joue Bernard Lubat sur un clavier de synthétiseur et à partir de ce qu'il a joué, on génère des séquences musicales. Le musicien joue sur la piste du haut où les notes sont représentées par les petits segments horizontaux (représentation dite « piano-roll » des données MIDI). Au milieu se trouve la piste de l'ordinateur qui est vide au début. En bas, on a des informations sur la pulsation qui est représentée par des petits points et des informations harmoniques puisque le logiciel va calculer des phrases selon la structure harmonique sous-jacente du morceau (ce qu'on appelle la *grille d'accords*). Les harmonies sont représentées par des petits traits horizontaux qui correspondent aux fondamentales des accords. Au début le musicien joue tout seul, puis il s'arrête et l'ordinateur prend le relai.





Le principe de ce calcul qui se base sur les traces laissées par le musicien (les notes qu'il a jouées et qui s'affichent sur le piano-roll) est devenu aujourd'hui le paradigme dominant dans toutes les applications d'intelligence artificielle qui se sont développées ces dernières années. C'est le cas notamment des algorithmes de recommandation qui connaissent un succès considérable sur les sites de commerce en ligne ou sur les plateformes de streaming vidéos ou musicales. Les traces laissées par les utilisateurs servent en quelque sorte à « prédire » le futur à partir du passé. Des applications aussi banales que les mécanismes de suggestion de clavier de nos téléphones portables mettent en œuvre un principe similaire (on propose à l'utilisateur de compléter les mots qu'il saisit en fonction des mots qu'il avait tapés antérieurement). Dominique Cardon est un sociologue spécialiste de l'Internet qui s'est intéressé à ces mécanismes de recommandation. L'une de leurs caractéristiques fondamentales est de ne faire aucune hypothèse sur la nature des traces laissées par les utilisateurs. Contrairement à une approche sociologique plus ancienne (par exemple celle de Pierre Bourdieu) qui se basait sur la définition de catégories socio-

professionnelles, l'approche mise en œuvre dans ces nouvelles techniques tend à considérer l'individu pour lui-même, sans référence à aucune catégorie préexistante. En ce qui concerne notre logiciel d'improvisation, la situation est similaire : le logiciel ne fait aucune hypothèse sur la nature de la musique jouée par le musicien dont le jeu est capté par la machine et aucune théorie musicale n'est intégrée *a priori* dans le système. C'est ce qui donne à notre logiciel des capacités d'adaptation très importantes. Nous l'avons utilisé dans toutes sortes de musiques de type jazz (Bernard Lubat) ou world (Charles Kely Zana-Rotsy) et nous faisons actuellement des expériences dans le champ de l'électro, mais rien ne s'oppose à ce qu'il soit utilisé dans d'autres champs, par exemple dans celui des musiques de la modalité arabe.

Le lecteur intéressé trouvera d'autres vidéos sur le site http://digitaljazz.fr qui montrent les développements récents de ce projet présentant une autre forme de connexion entre les mathématiques et la musique. L'un des écueils potentiels des techniques basées sur l'enregistrement du passé pour prédire le futur est d'enfermer l'usager dans les limites de ses propres routines. Comme le souligne Dominique Cardon dans son livre *À quoi rêvent les algorithmes ?*, elles le conduisent à « ajuster ses désirs sur la régularité de ses pratiques ». Plus généralement elles entérinent un certain ordre qui comporte toujours le risque de s'imposer comme une norme. Il est donc essentiel de laisser toujours ouverte la possibilité d'une étincelle créatrice qui échappe en quelque sorte à la régularité des traces enregistrées par la machine. C'est ce point qui rend nécessaire dans notre travail le prolongement du développement purement logiciel par une enquête de type socio-ethno-musicologique sur la manière dont les musiciens s'approprient un tel outil informatique. J'ai appelé *simulation* cette approche consistant à faire jouer le logiciel dans un certain contexte musical et à recueillir les réactions des musiciens sur les productions de la machine, et je renvoie le lecteur à mon article sur le sujet. D'une certaine manière, l'approche par simulation permet de relier le point de vue créatif et génératif adopté dans le développement de ce

logiciel avec le point de vue plus analytique adopté dans la première partie de cet exposé où les mathématiques étaient utilisées pour décrire des phénomènes musicaux. En simulant certaines pratiques musicales et en observant la manière dont les musiciens réagissent, on permet au logiciel d'improvisation de révéler comment la musique est organisée et ainsi de devenir un puissant outil d'analyse

Références :

✓ Amine Beyom, Systématique modale : modélisation des échelles musicales sur une grille de 24 quarts de ton, *Antouniah* n°6, Presses de l'université Antonine, Baabda-Liban, 2005, p. 228-236
http://foredofico.org/CERMAA/wp-content/uploads/2017/11/Beyhom_2006_Systematique_modale_UPA.pdf

✓ Srečko Brlek, Marc Chemillier & Christophe Reutenauer, Music and combinatorics on words: a historical survey, *Journal of Mathematics and Music, special issue on "Music and combinatorics on words"*, Volume 12, 2018 - Issue 3, p. 125-133
https://www.tandfonline.com/doi/full/10.1080/17459737.2018.1542055

✓ Dominique Cardon, *À quoi rêvent les algorithmes ? Nos vies à l'heure du big data*, Paris, Seuil, 2015.

✓ Marc Chemillier, trois notices pour le disque *La parole du fleuve. Harpes d'Afrique centrale*, Cité de la musique, CM001, 1999.
http://ehess.modelisationsavoirs.fr/marc/publi/disque/disque.html

✓ Marc Chemillier, Periodic musical sequences and Lyndon words, G. Assayag, V. Cafagna, M. Chemillier (eds.), *Formal Systems and Music, special issue of Soft Computing*, vol. 8(9), 2004, p. 611-616.

✓ Marc Chemillier, *Les Mathématiques naturelles*, Odile Jacob, 2007.

✓ Marc Chemillier et Gérard Assayag, OMax : présentation multimédia des recherches sur l'improvisation et l'ordinateur de l'Ircam et de la Cie Lubat, *Musimédiane*, numéro 3 *"Musiques non écrites"*, mai 2008
https://www.musimediane.com/numero3/chemillier/chemillier01.html
Animation multimédia :
http://ehess.modelisationsavoirs.fr/lubat/mediation/index.html

✓ Marc Chemillier, *De la musique aux mathématiques… et réciproquement*, vidéo pour le site VidéoDiMath, 2017 :
http://video.math.cnrs.fr/de-la-musique-aux-mathematiques-et-reciproquement/

✓ Marc Chemillier, De la simulation dans l'approche anthropologique des savoirs relevant de l'oralité : le cas de la musique traité avec le logiciel Djazz et le cas de la divination, *Transposition, Hors-série n°1, Musique, histoire, société. Les études sur la musique à l'EHESS*, 2018
https://journals.openedition.org/transposition/1685

✓ John Clough and Jack Douthett, Maximally Even Sets, *Journal of Music Theory*, Vol. 35, No. 1/2 (Spring – Autumn, 1991), p. 93-173.

✓ Godfried T. Toussaint, *The Geometry of Musical Rhythm,* Chapman and Hall/CRC, 2013.