

# Vers une sémiologie de la musique par modèles génératifs

## Le cas du contrepoint rigoureux

Guilhem Marion

16 juin 2017

### Résumé

Dans ce document nous allons essayer de considérer la musique comme un langage, et pour ce voir aller du côté de la linguistique afin de s'appropriier ses outils ainsi que de ses définitions. Nous nous intéresserons à la distinctions syntaxe-sémantique et montrerons leur dualité, puis nous ferons un tour des modélisations de langage possible et appliquerons la modélisation par contrainte au contrepoint rigoureux. Enfin par l'analyse du contrepoint au travers de ce langage, nous définirons la notion de *divergence de registre* et montrons la complexité de la dualité syntaxe-sémantique dans le cas de la musique.

# Table des matières

<b>Introduction</b>	<b>3</b>
<b>1 Contextualisation et définitions</b>	<b>3</b>
1.1 La musique est-elle un langage ? . . . . .	4
1.1.1 Le langage, un objet complexe . . . . .	4
1.1.2 Le cas de la musique . . . . .	8
1.2 Le Contrepoint rigoureux . . . . .	13
1.2.1 Histoire et principe . . . . .	13
1.2.2 Formalisation . . . . .	14
1.3 Modélisations grammaticales . . . . .	15
1.3.1 La dualité syntaxe/sémantique . . . . .	16
1.3.2 Noam Chomsky : grammaires génératives et langages formels . . . .	16
1.3.3 Modélisation de langages naturels . . . . .	19
1.3.4 Modélisation de langages musicaux . . . . .	19
<b>2 Méthode</b>	<b>21</b>
2.1 Modélisation par contraintes . . . . .	21
2.2 Construction du programme . . . . .	21
2.3 Test de la robustesse du modèle par test psychologique . . . . .	24
<b>3 Analyse et synthèse</b>	<b>27</b>
3.1 Résultats de l'expérience . . . . .	27
3.1.1 Les adjectifs dominants . . . . .	27
3.1.2 Les degrés de discrimination . . . . .	28
3.2 Analyse musicologique . . . . .	29
3.2.1 Analyse harmonique . . . . .	29
3.2.2 Le registre au sens littéraire . . . . .	32
3.2.3 Evolution de la <i>fonction sémantique</i> et concept de <i>connotation</i> . . .	33
<b>Conclusion</b>	<b>34</b>
<b>Bibliographie</b>	<b>35</b>
<b>Annexes</b>	<b>36</b>

# Introduction

Je considère la musique par essence impuissante à exprimer quoi que ce soit : un sentiment, une attitude, un état psychologique, un phénomène de la nature, etc. L'expression n'a jamais été la propriété immanente de la musique<sup>1</sup>.

Quoi de plus troublant que d'entendre pareille pensée ? La musique, impuissante à exprimer quoi que ce soit, que faire alors des frissons à l'écoute de *Lacrimosa*, le plaisir esthétique et intellectuel du Clavier bien tempéré, et même l'effroi et la grandeur du *sacre du printemps*. C'est sur Stravinsky est bien ingrat, mais si lui même peut dire chose pareille c'est qu'il bien une difficulté derrière cette question. Il est vrai que la musique met en émoi, fait pleurer, rire, danser ; la musique partage, dénonce, instruit et rassemble. Il paraît dès lors hasardeux de ne pas la considérer un langage ; mais alors qu'est-ce qu'un langage ?

Nous essaierons par cette question de tenter d'éclaircir cette citation, nous verrons pour cela un détour du côté de la linguistique et interrogerons Saussure, Sapir et Chomsky nous essaierons de comprendre ce qu'est un langage, et comment le caractériser. A cette fin nous nous intéresserons à la dualité syntaxe-sémantique dans les langages naturels et en particulier dans le français, puis, forts de ces informations, nous tenterons de les appliquer à la musique afin de voir si elle communique, ce qu'elle communique, et comment le sens en musique peut être produit. Nous aurons alors une idée claire de ce que peut être le langage musical.

Nous essaierons alors de décortiquer les méthodes de formalisation des langues pour entrevoir comment nous pourrions les utiliser en musique, nous nous attarderons particulièrement aux grammaires génératives de Chomsky et tenterons ce voir en quoi un modèle génératif, autant que tout autre modèle, si ce n'est plus, peut nous apprendre sur l'organisation de la musique en langage. Nous prendrons alors l'exemple du contrepoint et tenterons alors de répondre à la question : est-il envisageable de comprendre un langage musical par des méthodes utilisant des modèles génératifs ? Nous tenterons d'y répondre par l'expérience, en tentant d'éclaircir la question de la syntaxe-sémantique en musique à l'aide de ce modèle.

## 1 Contextualisation et définitions

Afin de mener à bien ce projet, il nous faut dans un premier temps bien définir notre objet musical, pour cela nous ferons un tour du contexte musical de son avènement, son

---

1. Igor Stravinsky, *Chroniques de ma vie*, Tome 1, Denoël, Paris, 1935 (réédition 2001)

évolution ainsi que son utilisation, puis nous le définirons brièvement à l'aide du précieux ouvrage de Marcel Bitsche *Traité de contrepoint* qui nous guidera tout notre travail.

## 1.1 La musique est-elle un langage ?

On entend souvent parler de langage en musique, que ce soit pour le langage tonal ou celui de Mozart ou Beethoven. Mais avant toute chose il faut se demander ce qu'est un langage et dans quelles mesures la musique peut se poser en langage ? Nous allons tenter de chercher en linguistique des définitions et pistes d'analyse choisies qui nous semblent cruciales à la compréhension de notre objet d'étude.

### 1.1.1 Le langage, un objet complexe

Malgré de nombreuses réflexions autour de la question du langage depuis l'antiquité, les linguistes s'accordent à définir son avènement par Saussure au début du XX<sup>ème</sup> siècle. Il définit le langage dans son cours de linguistique générale de la manière suivante :

Dans un premier temps, la langue<sup>2</sup>, fait initialement vocal et non écrit, ne peut se réduire au seul son. Les syllabes que nous manipulons sont des *impressions acoustiques* perçues par l'oreille et n'existeraient pas dans les organes vocaux. Le son dans la langue est déterminé par ces deux aspects qui sont indissociable, on ne peut donc étudier la langue par le seul aspect du son, ni par le seul aspect de l'articulation buccale. Dans toute étude de langage il est donc important de prendre en compte, le quoi et le comment de la production de phonème.

Mais est-ce le son qui fait le langage ? Selon Saussure, il est *l'instrument de la pensée* et n'existe pas pour lui-même, il est un moyen que trouve la pensée pour s'exprimer à autrui. Le son qui est déjà une unité complexe acoustico-vocale forme, avec la pensée, une autre unité complexe qui cette fois-ci est physiologique et mentale.

De ce fait la langue est double : elle possède un côté individuel et un côté social, et l'on ne peut concevoir l'un sans l'autre. À chaque instant elle implique à la fois un système établi et une évolution si bien qu'il est difficile de distinguer entre ce système et son histoire. Ces dualités font de la langue un objet extrêmement difficile à investir car si l'on s'attarde à un seul aspect du problème nous risquons de passer à côté de toute sa complexité, et si nous essayons de le comprendre dans son intégralité et nous apparaît comme un amas confus de choses hétéroclites.

Selon Saussure, il n'y a qu'une façon de contourner ce problème, ce n'est plus d'investir la langue, qui est la capacité d'utiliser un système de signes permettant à des individus de l'espace humain de communiquer, mais la parole, qui quant à elle est une instance du langage, un système de signes précis permettant effectivement à une communauté linguistique

---

2. Instance du langage, système de signes déterminé permettant à une communauté linguistique de communiquer.

de communiquer. Il nous faut donc étudier la langue comme manifestation du langage, car elle seule paraît être susceptible d'une définition autonome :

Mais qu'est-ce que la langue ? Pour nous elle ne se confond pas avec le langage : elle n'est qu'une partie déterminée, essentielle, il est vrai. C'est à la fois un produit social de la faculté du langage et un ensemble de conventions nécessaires, adoptées par le corps social pour permettre l'exercice de cette faculté chez les individus. Pris dans son tout, la langue est multiforme et hétéroclite ; à cheval sur plusieurs domaines, à la fois physique, physiologique et psychique, il appartient encore au domaine individuel et au domaine social ; il ne se laisse classer dans aucune catégorie des faits humains parce qu'on ne sait comment dégager son unité. La langue, au contraire, est un tout en soi et un principe de classification. Dès que nous lui donnons la première place parmi les faits de langage, nous introduisons un ordre naturel dans un ensemble qui ne se prête à aucune autre classification.<sup>3</sup>

Ainsi il apparaît que la langue serait une capacité humaine "innée", la capacité de communiquer par un système de signe, alors que la langue quant à elle est acquise et conventionnelle. Whitney<sup>4</sup> semble très juste sur ce point : la langue est une convention, et la nature du signe dont on est convenu est indifférente, arbitraire. La langue n'émane donc pas du sujet, elle est un produit social ; il nous faut donc faire la distinction entre langue et parole, qui quant à elle est un acte individuel de volonté et d'intelligence dans lequel on peut reconnaître une façon d'agencer les signes et des mécanismes psycho-physique permettant de leur donner du sens à l'échelle sociale, ces paramètres sont singuliers et permettent de différencier les individus par leur façon d'utiliser la langue.

La langue telle que la linguistique l'investit est un système de signes exprimant des idées par les sons et les mots, mais elle dans son sens le plus large comparable à l'écriture, la langue des signes, langage mathématiques, etc. On peut ainsi concevoir une science qui étudie la vie des signes au sein de la vie sociale : Saussure la nomme sémiologie (du grec *sémeion*, signe). Elle nous apprendrait en quoi consistent les signes et quelles lois les régissent. Cette démarche fut très présente au XX<sup>ème</sup> siècle, que ce soit pour les langues, les images<sup>5</sup>, ou la musique<sup>6</sup>, qui sera précisément notre sujet dans ce document.

---

3. Ferdinand de Saussure, *Cours de linguistique générale*, Payot, Paris, 1965, p.25

4. William Dwight Whitney né en 1827 et mort en 1894) était un linguiste, philologue, lexicographe et orientaliste américain. Connue comme le précurseur de la linguistique, ses travaux sont beaucoup utilisés par Ferdinand de Saussure.

5. Roland Barthes (1915 - 1980) est un philosophe, critique littéraire et sémiologue français. Il a beaucoup travaillé sur les signes visuels, notamment pour une sémiologie de la photographie (*La chambre claire*), des images produites par les modernités (*Mythologies*), ou pour décrypter les images quotidiennes d'une civilisation extra-européenne (*L'empire des signes*).

6. Jean-Jacques Nattiez, *Fondements d'une sémiologie de la musique*, Union Générale d'Édition, 1975, 448 p.

Edward Sapir<sup>7</sup> nous donne une définition simple, intuitive et concise :

Language is a purely human and non-instinctive method of communicating ideas, emotions, and desires by means of a system of voluntarily produced symbols.<sup>8 9</sup>

Par conséquent et plus généralement il y a langage quand deux individus peuvent communiquer, peu importe avec quel degré de précision, à l'aide d'un système de signe dédié. Il peut être question de communication par des sons, des signes visuels, des images, etc.

Plus formellement un langage peut se définir par un ensemble de *monème*<sup>10</sup> on peut, pour simplifier la chose, se référer aux dictionnaires contenant un ensemble de signes - qui sont définis par un signifiant<sup>11</sup>. et un signifié<sup>12</sup> lui conférant ainsi sa faculté de production de sens.

Pour qu'un langage soit compréhensible sur cet ensemble de signes, il faut qu'il soit régi par des règles ; elles seront de deux natures distinctes. Les règles syntaxiques assurent la cohérence du langage et la construction des phrases bien formées conformément à ce langage. Elle permet de construire des phrases grammaticalement correctes dans une langue donnée qui pourrait alors être interprétées par un locuteur de celle-ci. Elle ne peut cependant servir d'outil expressif ni faire référence à des états mentaux de conscience et d'intentionnalité tant que la *sémantique* n'est pas maîtrisée par tous les interlocuteurs. La sémantique correspond aux règles permettant de donner du sens à l'agencement des monèmes supposant le respect des règles syntaxiques ; sans syntaxe il ne peut y avoir de sémantique car les phrases ne peuvent même pas dégager de sens. Un langage ainsi défini avec ces conditions minimales permet de communiquer entre individus si elles sont définies au préalable et partagées par un ensemble d'individus souhaitant communiquer. Là est bien le premier problème de l'avènement du langage : comment définir et se mettre d'accord sur un ensemble de règles définissant un langage sans pouvoir communiquer par l'intermédiaire d'un langage préexistant. C'est une question fondamentale de la linguistique évolutive, que nous nous efforcerons de ne pas traiter.

Essayons désormais d'appréhender cette définition par un exemple issu de la langue

---

7. Edward Sapir (26 janvier 1884 – 4 février 1939) est un linguiste et anthropologue américain d'origine lituanienne. Il est considéré comme une figure des plus importantes en linguistique, son travail porta notamment sur l'acceptation du langage comme fait entièrement culturel (s'opposant ainsi à Chomsky) et par conséquent comme objet d'étude de l'ethnographie : la culture et la langue doivent être étudiées de pair.

8. Le langage est un moyen de communication purement humain et non instinctif pour les idées, les émotions et les désirs, par l'intermédiaire d'un système de symboles sciemment créés.

9. Edward Sapir, *Language, a introduction to the Study of Speech*, Harcourt, A harvest book, Brace and World, Inc., New York, 1949, p.8

10. Ou morphème selon les auteurs, qui correspond à la plus petite entité de sens, il se réduit souvent au mot dans la langue mais pas toujours, un mot peut évoquer plusieurs unités de sens, quand il est conjugué par exemple.

11. La représentation mentale du signe en lui même, qu'il soit sonore ou visuel

12. La représentation mentale de l'idée, concept ou objet auquel ce signe fait référence.

française. Le cas le plus simple est la phrase sous la forme

*Sujet · Verbe · Complément*

La syntaxe nous oblige à faire entrer chaque mots de notre dictionnaire dans une catégorie bien définie et à construire les phrases avec la formule suivante. Notons qu'il est pour l'instant difficile de définir formellement ces règles, c'est pourquoi nous consacrerons p.9 une partie consacrée à la définition de la notion des grammaire générative permettant de définir formellement des règles syntaxiques. :

$S \cdot V \cdot C$

**S** est le sujet, décrit comme un nom précédé d'un déterminant lui correspondant, accordé en genre et en nombre

**V** est le verbe accordé en genre et en nombre avec le sujet le précédant. S'il est transitif il pourra accepter un complément

**C** est le complément, sa nature dépend de celle du verbe : adjectif ou nom dans la plupart des cas, selon si le verbe est un verbe d'état ou non.

A l'aide de cette formule il est désormais aisé de construire des phrases bien formées sans même se soucier de son sens, on peut ainsi construire les phrases suivantes :

- Le chat mange la souris
- Le chien est gentil
- La chat mange la vie
- La lune est gentille

Malgré la syntaxe tout à fait correcte de ces phrases, elle ne sont pas toutes écrites dans un français correct car un sens ne peut pas toujours s'en dégager c'est là qu'apparaît le caractère nécessaire de la sémantique. Elle va nous permettre de donner du sens à l'agencement des mots. On définit alors le sens que peut prendre chacun des mots (nous nous épargnerons ici de rappeler ce qu'est un chat ou un chien) et le sens qui s'en dégage lors de leur agencement. Dans notre cas une règle sémantique serait de dire que le sujet est toujours l'acteur de l'action et que dans le cas de verbe transitif l'action s'effectue sur le complément, dans le cas d'un verbe descriptif on propose une propriété donnée par le complément au sujet. Ainsi, les phrases *Le chat mange la vie* et *La lune est gentille* ne peuvent plus être correctes car elles ne renvoient à aucune réalité (à part métaphorique ou poétique mais cela serait défini par le contexte). Malheureusement, la sémantique de la langue française est ambiguë, il convient donc d'éviter les cas d'ambiguïtés comme :

J'observe mon voisin au télescope

Il est en effet impossible dans ce cas de dégager du sens de la phrase car deux interprétations sont possibles :

1. (J'observe l'homme) au télescope.
2. J'observe (l'homme au télescope).

On ne peut donc savoir si c'est moi qui ait un télescope ou mon voisin. Pour lever l'ambiguïté, il faudrait ici avoir recourt à un système de parenthésage comme dans la notation mathématique.

La notion de non-ambiguïté n'est donc pas une condition nécessaire à la définition d'un langage.

### 1.1.2 Le cas de la musique

Maintenant que nous avons compris dans une certaine mesure ce qu'est un langage, qu'en est-il de la musique? Nous allons dans cette partie essayer de d'appréhender les notions vues plus haut aux prismes de la musique, de son potentiel expressif, et voir si elle peut satisfaire aux conditions nécessaires et suffisantes de la notion de langage.

Une chose essentielle est que d'après Edward Sapir un langage est avant un fait social qui lie les hommes entre eux, quelques en soit les signes et principes utilisés; et c'est précisément le cas de la musique qui est, de toute évidence un système de liaison entre les hommes et sa puissance fédératrice est immense. Il n'est plus à démontrer la puissance sociale est politique de la musique, qu'elle soit un liant d'une communauté (Blues, Raggaï), une force émancipatrice (Punk, Techno, etc.), et un acteur politique de taille, et ce depuis toujours (De Pierre-Jean de Béranger à NTM). Il est donc immédiat que la musique permet de communiquer de transmettre par agencement de signes musicaux qu'ils soient sous forme de notes de musique dans la musique tonale ou de sons plus abstraits en musiques Électroacoustiques. Peut-on parler pour autant de langage? Pas encore une système agencable ne suffit à un langage, il n'en est que la structure, les prémices. Il nous faut continuer notre réflexion afin de montrer que la musique peut mettre en place une syntaxe, une sémantique, et plus généralement ce quelle est capable de communiquer.

## Syntaxe

Pour commencer imaginons une petite expérience : si nous écoutons un choral de Bach, un concerto de Mozart et un opéra de Wagner nous serons capable de les identifier directement. Puis généralement, après avoir écouter un certain nombre de chorals de Bach nous serons capable de reconnaître un choral que nous n'avons jamais écouter de n'importe quel autre pièce musicale. C'est bien chacun des ces ensembles de pièces ont des propriétés communes qui diffèrent des autres ensembles de pièces. Imaginons maintenant que je prenne un choral de Bach et y modifie un certain nombre de choses, il y alors de fortes chances que vous le reconnaissiez pas que ça en est un. Cela veut dire une chose simple : nous identifions des genres et styles musicaux aux travers des propriétés que l'ensemble des pièces partagent. Ces propriétés représentent la façon dont sont construites ces pièces et s'expriment souvent sous forme de règles (c'est le cas des premiers traités d'harmonie),



bien entendu ces règles different d'une époque ou d'un genre à l'autre, on parlera alors de langues différentes ; elles different aussi entre compositeurs, la notion de parole vue plus haut entre alors en compte. Nous appellerons ces règles structurelles permettant d'identifier une pièce musicale la syntaxe. C'est en effet la façon d'agencer des signes musicaux pour leur donner une cohérence qui leur permettra alors d'être interprétés par l'auditeur. Dans certains cas la syntaxe d'un langage est formellement définie et peut même l'être à partir d'arguments théologique, philosophique ou mystique, loin de toute réalité musicale. Dans d'autres cas, la syntaxe peut se construire empiriquement, le rôle du musicien sera alors de la comprendre, de l'intérioriser par l'écoute et la reproduction ; c'est notamment le cas des musiques improvisées, dont le Jazz où une formalisation est presque impossible : c'est par l'essai-erreur, l'écoute et la pratique que l'on devient musicien de Jazz. Ainsi pour rentrer dans un cadre ou une communauté musicale, faire partie d'un courant, il faut s'imposer les règles, les apprendre, les pratiquer, ce qui n'empêche pas bien sur d'y apporter sa sensibilité c'est le but du jeu : de s'exprimer avec des contraintes dans un langage intelligible. Dans ce cas, que faire des créateurs ? Des compositeurs pour qui le seul objectif et justement de creuser là où personne n'a jamais creuser, d'inventer ce que personne n'a jamais inventer, de créer et jamais reproduire ? Ces personnes là font bien entendu partis du processus, ils sont les créateurs de règles, dans leur travail même ils créent un langage, leur langage. Ils sont les moteurs de l'évolution des langues musicales, que ce soit dans leur champ d'application (un compositeur d'opéra voulant innover sur tel ou tel aspect faire évoluer le genre le l'opéra) ou dans le langage musical (l'avènement de la musique électronique permet aujourd'hui de repenser complètement absolument tous les styles musicaux de la pop au Jazz, et cela par le biais de l'utilisation de ces outils par les premiers compositeurs électroniques).

Les choses sont désormais plus claires, on appellera *langage musical* la capacité que les hommes ont d'organiser des sons musicaux<sup>13</sup> afin de communiquer entre individus. Une *langue musicale*, quant à elle, est une instance d'un langage, c'est une réalité musicale régie par un ensemble de règles ou propriétés communes que l'on définira de syntaxiques. La *parole musicale* enfin est l'expression personnelle d'un individu à travers un langage donné. Nous avons aperçu comment pouvait naître une langue et quels pouvaient être les moteurs de son évolution. Cependant si l'on parle de langue et de syntaxe il faut parler de sémantique et de sens, comment la syntaxe s'organise pour créer du sens et de quel nature ce sens peut-il être ?

## Sens

Comme nous l'avons vu plus haut la sémantique correspond à des règles qui permettent de tirer du sens de l'organisation de signes, c'est un système interprétatif. Ce système comme nous l'avons vu pour la langue française n'a pas besoins d'être non-ambiguë c'est

---

13. La notion de sons musicaux s'élargit aujourd'hui comme tout son utilisé dans un contexte musical, il peut s'agir en électroacoustique de n'importe quel son.

à l'individu de jouer avec ou d'éviter ces ambiguïtés. Il faut néanmoins remarquer deux choses.

La première est que le langage musical contrairement aux langues dans leur sens initial ne possède pas des signes de même nature : un signe pour Saussure est défini par la dualité signifiant/signifié ce qui en fait le plus souvent un monème. Les signes musicaux (typiquement les accords en musique tonale) n'ont pas, à proprement parlé, de référent<sup>14</sup> mais plutôt une couleur, un sentiment mais qui peut sensiblement changer selon son contexte d'utilisation. C'est donc principalement l'agencement de ces signes qui crée du sens et non l'agence de plus petites entités de sens. Mais ce point sera vite éclairci.

Le second point est que du fait de la singularité, la subjectivité et l'évolution rapide de l'utilisation du langage musical on ne peut en évaluer le sens absolu d'un énoncé. C'est un point auquel Jean Molino et Jean-Jacques Nattiez apportent une importance toute particulière<sup>15</sup>.

### Tripartition de Molino

Si l'analyse sémiologique est l'entreprise qui prend en charge la manière dont la musique devient *fait symbolique* pour ses utilisateurs, il faut bien spécifier lesquels car ils sont multiples : le compositeur, l'interprète, le musicologue, l'auditeur. Comment travailler avec cette spécificité ? C'est à ce point que Jean-Jacques Nattiez fait référence dans *Fondements d'une sémiologie de la musique* à la conception sémiologique de Molino selon lequel le processus de symbolisation implique trois pôles : le message en lui-même dans sa réalité matérielle, les stratégies de production du message et ses stratégies de réception. L'analyse d'une oeuvre musicale comme fait social, de part sa nature langagière, doit impérativement investir le contexte de création et réception d'une oeuvre et tenter de comprendre différentes interprétations car la musique comme tout langage peut être mésinterprétée. D'une part à cause d'éventuelles définitions flottantes quant à sa sémantique (comme nous l'avons vu plus haut avec le cas de la langue française) mais aussi à cause de son évolution perpétuelle ainsi que des multiples processus subjectifs intervenant de sa création à sa réception.

Selon Molino<sup>16</sup> la description d'un message en lui-même est *l'analyse du niveau neutre*, la *poiétique* correspond aux processus de création, et l'*esthétique* aux processus de réception. Le niveau neutre est l'étape intermédiaire entre le créateur et l'auditeur elle se situe précisément entre la volonté significatrice du compositeur et le sens que peut l'auditeur peut en tirer. Le niveau neutre correspond donc à l'oeuvre à nue, sans aucune signification

---

14. Défini par Saussure comme l'objet ou l'ensemble des objets désigné par le signe

15. Jean-Jacques Nattiez, *Fondements d'une sémiologie de la musique*, La tripartition de Molino et spécificité du symbolique, Union Générale d'Édition, 1975, p.50

16. Initialement cette distinction tient de la branche phonologique de la linguistique : la phonétique se divise en une phonétique articulatoire (poiétique), une phonétique acoustique (niveau neutre) et une phonétique auditive (esthétique). Ces disciplines s'est refusée à considérer la description physique des sons comme correspondant à la manière dont ils sont produits et perçus.

propre. Il correspond à un ensemble de signes (continuons la rapprochement avec les langages parlés) que l'on essaierait d'interpréter sans se soucier d'un système sémantique qui lui conférerait un sens.

#### Shéma

Pour parler de sémantique en musique il faut donc être très attentif à son contexte : parle-t-on du sens que l'auteur veut conférer à l'oeuvre ou que nous (ou quelqu'un d'autre) y donne par le biais de son interprétation et de sa sensibilité. Si l'on veut espérer parler du niveau neutre il faut s'assurer d'une éventuelle convergence des niveaux poétique et esthétique, d'une convention - même locale, ambiguë ou restreinte à une petite communauté d'individus - ce qui confèrera une dimension langagière à l'énoncé musical. Si cette condition n'est pas du tout respectée (car elle ne peut jamais l'être entièrement), la musique perd ainsi son statut de langage, elle devient une curiosité, objet esthétique venu d'ailleurs que l'on utilise, sorti de son contexte, dans l'objectif d'une expérience esthétique nouvelle. Comment, en effet, parler de message si l'on a pas les clefs pour le déchiffrer, dans ce cas il n'y ni message, ni langage, et encore moins sémiotique. Cette distinction est bien entendu une distinction pratique prenant en compte le caractère dynamique et subjectif de la musique ; le concept de *compétence* que l'on retrouve chez Chomsky confond quand à lui la poétique et l'esthétique. La compétence est à la fois la faculté de produire et d'émettre des énoncés chargés de sens qui n'ont jamais été dis auparavant. Cette combinaison conceptuelle marche très dans la théorie, quand étudie un langage comme un objet statique, figé et formellement défini, ce sera le cas de l'approche Chomskienne des langages.

Il est convenu du dire que la musique exprime des sentiments, des affects - ceux du compositeur en particulier<sup>17</sup>. Il paraît cependant plus juste de dire que la musique *engendre* des sentiments, et c'est précisément sur cette distinction que la tripartition de Molino est extrêmement pertinente. De quelle façon qu'elle soit composée, jouée, diffusée la musique engendre des affects subjectifs. La musique de fait, de part sa nature provoque à l'auditeur une production de sens, un sens personnel, de l'ordre de l'esthétique. Ainsi la musique fait sens, par essence. Et c'est ainsi que l'on peut apprécier une composition de musique africaine et arménienne sans pour autant *comprendre* les éventuelles significations de celle-ci ni les intentions de son compositeur, nous lui donnons du sens sans le vouloir, simplement par l'écoute. On peut ainsi trouver très joyeux une marche funèbre ou très angoissante une musique de danse. La musique n'est pas strictement dans ce cas un moyen de communiquer entre les hommes mais un moyen d'engendrer du sens, de créer du sens ensemble. De plus la genèse d'une oeuvre nous échappera toujours, du plus que nous ne n'en connaissons pas l'auteur. Malgré toutes les études biographiques, historiques, psychologiques que nous pouvons faire nous ne pourrons jamais déceler l'intention pure du créateur, l'oeuvre seule est la trace de cet acte, et nous ne pouvons faire autrement que la recevoir à travers notre sensibilité, l'essence de l'oeuvre nous est toujours inconnue et c'est peut être précisément cet invisible qui fascine et rend la musique si belle. Il est à préciser qu'il en est de même

---

17. *Le langage musical*, André Boucourechliev, Fayard, Les chemins de la musique, 1993

du langage parlé, on ne se comprend pas toujours et une idée, un énoncé n'est souvent pas interprété comme son producteur le voudrait ; mais il faut toujours sens chez le récepteur.

## Sémantique

La sémantique, à proprement parlé, ne peut alors se trouver que sous une forme locale, et peut prendre corps par des procédés précis ou au sein d'une communauté restreinte d'individus. Quand nous évoquons le langage nous parlons de communication et de conventions. C'est précisément ici que la sémantique apparaît, comment la musique peut servir communiquer, se comprendre ? Cela implique une convention, l'établissement d'un système sémantique permettant de tirer du sens d'énoncés musicaux, énoncés musicaux qui, bien entendu, peuvent prendre un sens tout à fait différents selon le système sémantique par lequel ils sont interprétés<sup>18</sup>. On peut néanmoins se fixer, dans une communauté définie à un moment défini, une sémantique une façon de tirer du sens d'un agencement de sons musicaux- que l'on pourrait appeler procédé musical. C'est le cas très précisément des théories sur le beau et le laid en musique, il est admis à l'époque baroque que les quintes et octaves parallèles provoquent chez l'auditeur une sensation désagréable, il n'est donc pas moins vrai qu'une pièce en possédant reste une pièce de musique baroque (Bach utilise dans ses chorals de nombreuses reprises des quintes parallèles ainsi que des fausses relations). Pour cela, une communauté s'est accordée sur le fait que cet agencement de notes en particulier évoquait un sentiment particulier ; c'est un début de système sémantique. Il est à préciser que contrairement à ce que l'on peut lire chez certains théoriciens de l'époque, ces considérations sur le beau et le laid sont loin d'être universelles ; d'autres cultures et même le jazz et les musiques contemporaines utilisent à souhait ces procédés considérés à l'époque comme impurs. On est alors face à une convention, un accord au sein d'une communauté précise à une époque donnée, permettant de donner du sens à des procédés musicaux. Ce système sémantique peut aller beaucoup plus loin que la simple considération du beau et du laid. Un catalogue de figures-formules a même été dressé au XVIII<sup>e</sup> siècle permettant de donner le sens des formules les plus usitées.

Parmi ces figures on peut citer, à titre d'exemple, le chromatisme descendant qui connote indifféremment la tristesse, le regret et le deuil ; ou inversement, les figures ascendantes par intervalles plus larges, rythmiquement élancées vers la note la plus aiguë qui connoteront au choix l'héroïsme, la volonté, la joie, la fierté<sup>19</sup>.

Souvenez-vous aussi de l'époque avant l'avènement du tempérament égal, l'époque de fait de l'inégalité des intervalles, toutes les tonalités avaient leur saveur propre, singulière. On retrouvait alors selon les lieux et les époques des grilles d'interprétation de chacune des tonalités à un affect particulier, les compositeurs les utilisaient alors dans ce sens. Par

---

18. Une phrase d'une langue a une autre peut avoir des significations totalement différentes et sont souvent sujet d'amusements ou de malentendus.

19. Ibid. p.11

exemple la tonalité de do mineur était pour Charpentier «Obscur et triste»<sup>20</sup>, mi mineur pour Rameau convenait à la douceur et à la tendresse<sup>21</sup>, Fa Majeur quant à lui évoquer la complaisance et le repos pour Schubart<sup>22</sup>, sans oublier Mozart pour qui la tonalité de ré mineur était systématiquement utilisée pour les requiem et l'évocation de la mort et du deuil. L'utilisation des tonalités a ainsi pris une place conséquente en musicologie, elle permet alors de dégager du sens de ces procédés et ainsi de comprendre les intentions du compositeur par le biais d'un système sémantique.

Nous avons ici montré que la musique peut se poser en langage, elle peut se présenter comme un outil de communication entre les hommes, de production de sens, pouvant posséder une syntaxe ainsi qu'une sémantique, mais pouvant aussi prendre ses libertés pour inventer, innover et surprendre. Elle n'a alors rien à envier aux langues parlées et écrites. Pour continuer notre réflexion nous allons essayer de voir comment la considération de la musique en langage peut nous aider à la comprendre, et pour ça, après de multiples considérations théorique, rien de vaut un exemple concret. Pour cela nous allons tenter d'investir le contrepoint rigoureux et tenter de vérifier ce qui a été dit plus haut, et d'appliquer notre raisonnement à cet objet tant étudié de tant de façons diverses et variées.

## 1.2 Le Contrepoint rigoureux

### 1.2.1 Histoire et principe

Afin de comprendre le contrepoint, il faut avant tout se replacer dans le contexte musical des premiers répertoires polyphoniques du 9ème siècle ; la polyphonie ne représentait pas encore de accords s'enchainant et ayant chacun une fonction harmonique classable comme l'harmonie le propose à partir du 17ème siècle. Une polyphonie se composait par superpositions successives de voix mélodiques, l'harmonie, dès lors, ne concerne pas les accords mais plutôt les intervalles, les rencontres des fois n'ont pas de fonctions dans la phrase musicale, mais doivent avoir de bons rapports en elles. Les premiers contrepoints chrétiens sont improvisés, on les appelle des *organa*, naissent de la volonté d'enrichir le langage musical liturgique, ces improvisations se construisent à partir d'une voix principale issue du répertoire grégorien, l'autre voix est improvisée. L'organum est le principal genre polyphonique du IX au XIIIème siècle, puis fut complété par le motet (donner + d'info). Il existe trois façon de composer un organum : il est parallèle si la voix ajoutée est une doublure à la quarte ou à la quinte de la voix principale, il est dit composite si la doublure est à l'octave, il est modifié si la voix part de l'unisson et y revient après un parcours mélodique. Ces premiers contrepoint étant improvisé, les seules données nous permettant de les comprendre sont les traités datant du IXème siècle et rien ne peut nous garantir leur application en pratique. Dès le début du XIème siècle un procédé va faire

20. *Règles de composition*, M.A. Charpentier (1636-1704), Paris, 1690.

21. *Traité de l'Harmonie*, J.P. Rameau, chap. 24, livre second, Paris, 1722

22. *Ideen zu einer Ästhetik der Tonkunst*, Chr. Fr. D. Schubart, Vienne, 1806

faire un grand pas en avant au contrepoint, c'est le *déchant*<sup>23</sup>. Il s'agit de faire évoluer une seconde voix au dessus du plain chant mais cette fois-ci en mouvement contraire ce qui permet d'éviter tout mouvement parallèle, de faire entendre un nombre considérable d'intervalle lors d'une même pièce et de donner une autonomie considérable à la voix organale. On compose désormais la musique note contre note, ce qui explique l'appellation de contrepoint. Plus tard la voix principale réalise le thème grégorien en valeur longue, ce qui permet à la seconde voix de prendre plus de liberté et de lui laisser le temps de se développer. C'est ainsi à cette époque que l'écriture musicale prend une place toute particulière de part l'importance de noter pour fixer ces polyphonie et pouvoir les composer. Au Xème et XIème siècle, à l'époque de l'école de Saint-Martial de Limoges, on retrouve des contrepoints écrits qui peuvent être de complexité rythmiques remarquable et les *organa Viderunt omnes et Sederunt principes* de Pérotin inaugure la musique à quatre voix.

Très vite les compositeurs ont rechercher les effets procurés par l'enchaînement des rencontres entre les différentes voix, on parle de consonance et de dissonances qui sont utilisées en alternance sur les temps forts et les temps faibles. La classification des intervalles selon ces deux notions est étonnamment complexe et diffère selon les auteurs et les époques. A la renaissance, le contrepoint s'est considérablement enrichi, donnant naissance à un riche univers rythmique, si bien que dès le XIVème siècle les quintes parallèles sont évitées et les premières cadences voix le jours. Une cadence étant définie comme une fin de phrase ou une petite respiration arrivant sur une consonance parfaite par mouvement contraire et conjoint ; dès lors l'harmonie tonale commence à germer.

### 1.2.2 Formalisation

Toute notre étude se base sur le traité de contrepoint Noel-Gallon et Marcel Bitsch<sup>24</sup>, il propose d'enseigner le contrepoint par un ensemble de 122 points allant de de l'écriture à deux voix à la fugue ; ces règles garantisse la production de contrepoint valides d'un point de vu purement syntaxique mais en aucun cas qu'une quelconque beauté ou idée esthétique puisse s'en dégager, c'est le travail de l'élève compositeur d'utiliser ces règles pour s'exprimer pleinement.

Nous nous restreindrons à l'écriture fleurie à deux voix qui survole une grande partie de l'ouvrage. Nous allons tenter ici de synthétiser quelques unes de ces règles, dont les plus importantes à nos yeux, afin de bien définir notre objet d'étude et de montrer sa manipulation par des modèles formels comme les grammaires de Chomsky.

Le contrepoint est la science des lignes mélodiques et de leurs superpositions ; il considère la musique sous son aspect horizontal, il est complémentaire de l'harmonie. Le contrepoint rigoureux ou scolaire enseigne à superposer de courtes mélodies vocales non modulantes ; ces mélodies doivent être composées et combinées à un cantus firmus suivant des règles

---

23. Micrologus, Guido d'Arezzo, 1026.

24. *Traité de contrepoint*, NOËL-GALLON, MARCEL BITSCH, Durand et Cie, 1964, 130 p.

strictes. Le cantus firmus est la mélodie imposée, écrites en valeurs longues. Le contrepoint peut s'écrire dans les deux modes classiques mais aussi dans les modes anciens.

## Notions fondamentales

Les voix se combinent au cantus firmus et se combinent entre elles de façon à se retrouver toujours en consonance au premier temps de chaque mesure. Entre ces points d'appui, une certaine liberté est laissée à chaque voix. De cette liberté naissent des rencontres de notes inattendues qui font tout l'intérêt du contrepoint. L'entrée d'une voix se fait forcément sur une consonance.<sup>25</sup>

Les intervalles admis comme consonants sont les suivants<sup>26</sup> :

- L'octave juste
- La quinte juste
- La tierce diatonique (mineure ou majeure)
- La sixte diatonique (majeure ou mineure)
- L'unisson est accepté sur un temps faible ou sur la partie faible du temps, il est permis au temps fort sur les première et dernière mesures. L'arrivée sur l'unisson est excellente par mouvement contraire ainsi que par mouvement oblique<sup>27</sup>.
- La quarte juste, la quarte augmentée et la quinte diminuée sont permises en notes mélodiques ainsi qu'entre deux parties ne comprenant pas la basse<sup>28</sup>.

Seuls les accords parfait majeur et mineur et leur premier renversement sont autorisés ainsi que l'accord de sixte, premier renversement de l'accord de quinte diminuée<sup>29</sup>.

Le mouvement conjoint doit être employé le plus souvent possible et poursuivi le plus loin possible, avec le mouvement contraire il constitue l'essence même du contrepoint<sup>30</sup>. Il faut particulièrement éviter les mouvements contraires à la barre de mesure<sup>31</sup>.

Les intervalles mélodiques permis sont tous les intervalles diatoniques majeurs ou mineurs ne dépassant pas la sixte et l'octave juste, l'intervalle chromatique, les intervalles diminués ou augmentés sont interdits<sup>32</sup>.

La quarte augmentée franchie en trois ou quatre notes doit être précédée ou suivie d'un mouvement conjoint de même direction<sup>33</sup>.

On ne peut arriver par mouvement direct sur des tierces, quarts ou sixtes.<sup>34</sup>

---

25. Ibid. § 5

26. Ibid. § 7

27. Ibid. §53

28. Ibid. §54

29. Ibid. §8

30. Ibid. §24

31. Ibid. §25

32. Ibid. §27

33. Ibid. §28

34. Ibid. §41

Maintenant que nous voyons plus précisément ce que représente le contrepoint ainsi que les règles qui le définissent, essayons de voir comment il serait possible de modéliser à la manière d'un langage et ce qui serait possible de modéliser.

### 1.3 Modélisations grammaticales

Depuis les travaux d'Alan Turing sur l'automatisation du traitement des données, l'idée de modéliser des objets complexes par traitement automatique s'est énormément répandue. Elle prend une place très importante pour définir des grammaires, que ce soit dans le cadre des langages formels en informatique théorique qu'en linguistique. Nous avons vu que la musique et plus particulièrement le contrepoint est propice, de part sa nature de langage formellement défini, à une modélisation. Nous essaierons dans un premier temps de comprendre ce qui peut être modéliser dans un langage, et comment cela peut se manifester.

#### 1.3.1 La dualité syntaxe/sémantique

A partir des années 1950, un nouveau champ disciplinaire apparaît : la linguistique informatique. C'est un champ interdisciplinaire composé entre autres de la linguistique et de l'informatique. Un des aspects de cette discipline est la recherche d'une formalisation du fonctionnement du langage humain permettant sa reproduction par un système artificiel. On distingue deux approches : statistique et logico-grammaticale. Mais dans quelle mesure est-il possible de modéliser des compétences langagières ? Pour cela intéressons nous à une expérience de pensée du philosophe américain de l'esprit, John Searl : la chambre chinoise. Dans cette expérience de pensée, Searle imagine une personne n'ayant aucune connaissance du chinois enfermée dans une chambre. Dans cette chambre, il y a un catalogue contenant un ensemble de règles permettant de construire des phrases correctes en chinois. Ces règles sont parfaitement claires et reposent uniquement sur la syntaxe des phrases. Un véritable sinophone situé à l'extérieur de la chambre pose des questions à la personne enfermée qui est ainsi capable d'y répondre en utilisant le catalogue. Du point de vue du sinophone, la personne dans la chambre se comporte comme un individu qui parlerait vraiment le chinois. Cependant, elle ne fait qu'appliquer les règles du catalogue et n'a aucune compréhension des phrases qu'elle entend et produit.

Ce test apparaît dans les années 1980 pour tenter de savoir si un programme informatique peut accéder à la compréhension d'un langage. Il vise à montrer que seulement une intelligence faible<sup>35</sup> peut être atteinte et ne fait que simuler une compréhension d'un langage. Mais elle nous intéresse dans notre cas car elle nous permet de faire la distinction entre syntaxe et sémantique et nous dit qu'il est rationnellement possible de modéliser la syntaxe d'un langage mais qu'un programme ne peut réellement accéder à sa sémantique, il peut seulement la simuler<sup>36</sup>. C'est pourquoi dans ce travail nous allons nous intéresser

---

35. Une intelligence artificielle ne présentant pas d'états mentaux de conscience et d'intentionnalité.

36. Ce point fait débat, en particulier pour les cognitivistes qui pensent que la pensée est analogue à



uniquement à la syntaxe de langages musicaux, c'est à dire à la forme qu'ils doivent prendre pour être considérés comme corrects, ou relevant d'un style bien déterminé, le *sens* musical quant à lui est laissé de côté lors de cette modélisation qui est essentiellement formelle, tout comme le sont les modélisations de Chomsky.

### 1.3.2 Noam Chomsky : grammaires génératives et langages formels

Une partie des travaux de Noam Chomsky durant la seconde moitié du XX<sup>me</sup> reposait sur ces questionnements quant au langage naturel. Très intéressé par les travaux se faisant en calculabilité sur les automates et machines de Turing, il postule qu'un langage formel tel que définie en mathématiques et informatique théorique peut s'écrire sous la forme d'une grammaire générative décrivant des systèmes de règles de production capable d'engendrer toutes les phrases correctes d'un langage naturel, il a alors pour but de modéliser formellement la syntaxe d'un langage.

Avant de rentrer dans les aspects techniques des grammaires selon Chomsky voici une définition d'une grammaire de façon générale, cela nous aidera à comprendre la pertinence du formalisme mathématiques dans son raisonnement.

«The grammar of a language can be viewed as a theory of the structure of this language. Any scientific theory is based on a certain finite set of observations and, by establishing general laws stated in terms of certain hypothetical constructs, it attempts to account for these observations, to show how they are interrelated, and to predict an indefinite number of new phenomena.<sup>37</sup> »

Les grammaires génératives s'inscrivent dans le courant de la linguistique générative principalement présente en Amérique du Nord, elle s'est développée depuis 1957 sous l'impulsion de Noam Chomsky. Elle tendent de reproduire les mécanismes cognitifs permettant à un locuteur de produire un nombre infini d'énoncés corrects et ceux sans jamais les avoir entendus. Il postule l'existence d'une *grammaire générative* permettant de construire à partir de règles formellement définies une infinité d'énoncés. Son travail porte son observation non sur la *performance*<sup>38</sup> en tant que telle, mais sur les mécanismes de *compétence*<sup>39</sup>. Ainsi, elle tente d'expliquer les règles que le locuteur applique de façon intuitive. Pour cela Chomsky s'inspire largement des travaux déjà fait en langages formels et a abondamment recourt au formalisme mathématiques pour définir son modèle. Nous essaierons dans cette

---

un processus de traitement de l'information, théorie qui s'oppose au béhaviorisme. Les computationalistes (Hilary Putnam en particulier) vont plus loin en penser le cerveau comme une machine et la pensée à un calcul, ces raisonnements permettent d'ouvrir la voix à l'intelligence artificielle et de repenser la dualité syntaxe-sémantique dans le cas des modélisations de langages.

37. Noam Chomsky, *Three models for the description of language*, IRE Transactions on Information Theory ( Volume : 2, Issue : 3, September 1956 )

38. L'emploi effectif d'une langue par un individu dans des situations concrètes, c'est la partie sociale de la langue. Elle s'oppose à la compétence.

39. Notion qui nous avons déjà vu plus haut, elle correspond à la capacité d'un locuteur de comprendre et produire des énoncés.

partie d'être aussi clair que nous le pouvons en essayant de rester le plus fidèle possible à la pensée de Chomsky.

Il définit d'abord la notion de langage, c'est un ensemble (bien souvent infini) de mots composés de lettres sur un alphabet définit au préalable.

L'idée de relier l'ensemble des mots (qui représentent plutôt des énoncés selon notre définition) qui appartiennent au langage et la façon dont ils peuvent être produits par des règles locales de leur agencement.

Un langage  $L$  est un ensemble de mots  $\omega_i$  sur un alphabet  $\Sigma$ .

$L = \Sigma^*$  est l'ensemble des mots que l'on peut écrire avec l'alphabet  $\Sigma$ .

Par exemple,  $L = \Sigma^*$ , avec  $\Sigma = \{a, b\}$  est l'ensemble des mots contenant les lettres  $a$  et/ou  $b$ .

$L = \{a^n b^n c^n \mid n > 0\}$  est l'ensemble des mots contenant autant de  $a$  que de  $b$  et que de  $c$  et pouvant se diviser en trois séquences de même longueur dont une ne contenant que des  $a$  et l'autre que des  $b$  et la dernière que des  $c$ .

On remarque que tout langage  $L$  sur l'alphabet  $\Sigma$  est un sous ensemble de  $\Sigma^*$ ,  $\Sigma^*$  représentant l'ensemble des possibilités sur  $\Sigma$  qui lui même peut représenter n'importe quel alphabet, il est clair que pour tout langage naturel  $k$  il existe un langage  $L_k$  sur  $\Sigma_k$  tel que  $L_k$  représente toutes les phrases valides à partir de l'ensemble des mots de cette langue, ici  $\Sigma_k$ . On comprend ainsi mieux pourquoi Chomsky s'est intéressé à cette théorie dans l'espoir de mieux comprendre les langues.

Afin de tenter de comprendre les mécanismes cognitifs à l'oeuvre lors de la production des énoncés, il introduit la notion de grammaire formelle<sup>40</sup> et en explique la pertinence pour les langages naturels.

Une grammaire  $G$  se définit par :

1. Un ensemble fini  $N$  de symboles non terminaux
2. Un ensemble fini  $\Sigma$  de symboles terminaux
3. Un ensemble fini  $P$  de règles de construction de la forme :

$$(\Sigma \cup N)^* N (\Sigma \cup N)^* \rightarrow (\Sigma \cup N)^*$$

4. Un symbole de départ  $S \in N$  permettant d'engendrer la grammaire

Prenons un exemple simple : on veut représenter sous forme de grammaire la façon dont on peut écrire des phrases simples avec un sujet, un verbe et éventuellement un complément. C'est à dire de la forme Sujet Verbe ou Sujet Verbe Complément. La grammaire sera alors  $G$  telle que :

---

40. Noam Chomsky, *Three models for the description of language*, IRE Transactions on Information Theory ( Volume : 2, Issue : 3, September 1956 )

$N = \{S, V\}$ ,  $\Sigma = \{sujet, verbe, complément\}$ ,  $S$  est le symbole de départ et  $P$  consiste en les règles suivantes :

$$S \rightarrow sujet \cdot V^{41}$$

$$V \rightarrow verbe$$

$$V \rightarrow verbe \cdot complément$$

Analysons cela pas à pas.  $S$  étant le symbole initial pour générer une phrase il faut obligatoirement commencer par lui, ce qu'il engendre sera notre phrase. Or il faut que notre phrase ne contienne que des symboles terminaux, il faudra donc replacer tous les symboles non terminaux par des symboles terminaux. Dans notre cas  $S$  engendre  $sujet \cdot V$ , il faut donc voir ce que peut engendrer  $V$  on remarque alors que les deux seules possibilités sont :  $sujet \cdot verbe$  et  $sujet \cdot verbe \cdot complément$ , ce qui correspond à ce que nous voulions représenter.

Cet outil permet à la fois de se prouver qu'il est possible de représenter formellement la syntaxe d'un langage, d'essayer de comprendre la façon dont les énoncés sont produits par le locuteur et d'avoir la main sur tous les énoncés valides d'un langage donné. Il n'est bien entendu pas question ici d'exhiber la grammaire formelle représentant le contrepoint rigoureux, cela serait trop laborieux pour un intérêt trop faible. Nous allons par contre montrer plus loin que le contrepoint est représentable par une grammaire générative, puis construire informatiquement une grammaire l'engendrant par addition de règles locales.

### 1.3.3 Modélisation de langages naturels

La méthode que nous venant de voir permet de représenter un langage par un système de règles qui permet ensuite de générer tous les énoncés valides de ce langage, nous appellerons cette approche l'approche formelle car elle s'appuie sur la forme du langage et le définit formellement, on y opposera l'approche statistique qui elle modélise un langage empiriquement à partir d'un grand corpus.

Des exemples de modélisation de langages naturels sont tentative de représentation de la langue anglaise à partir d'une grammaire de Chomsky hors contexte<sup>42</sup> et de l'utilisation d'outils statistiques pour la modélisation de style d'écriture dont notamment le projet *What would Shakespeare say ?*<sup>43</sup> qui permet de comparer différentes instances de la langues anglaise.

---

41.  $\cdot$  est l'opérateur de concaténation cela signifie simplement que l'on met cote à cote les deux opérandes.

42. *Speech and Language Processing.*, Daniel Jurafsky James H. Martin. Chapter 11 : Formal Grammars of English. Draft of January 30, 2017.

43. Le site internet *What would Shakespeare have said ?* permet de prédire statistiquement la fin de phrases selon la "grammaire de Shakespeare". L'article *Natural language processing : What would Shakespeare say ?* du site Gigadom se propose de comparer ces résultats avec un programme similaire modélisant la grammaire anglaise actuelle.

### 1.3.4 Modélisation de langages musicaux

Afin de bien comprendre en quoi consiste ce modèle statistique et dans quelles mesures la modélisation grammaticale peut s'appliquer à la musique, donnons un exemple d'une modélisation statistique de différentes grammaire de Jazz.

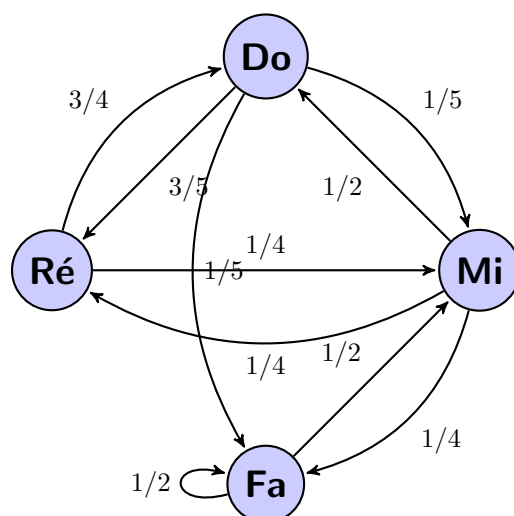
Il est en effet très rare que des langages comme le contrepoint soient aussi bien définis, dans le cas contraire il faut trouver une autre façon d'établir ces règles. Une méthode très courante est l'apprentissage statistique. Elle consiste à étudier un large corpus d'un langage, un programme en définit l'alphabet et calcule la probabilité d'apparition de chacune des lettres en fonction des  $n$  précédentes : on parle de modèle de  $n$  gramme qui correspond à un modèle de Markov d'ordre  $n$ . Si l'on applique ce modèle au langage musical, on modélise la syntaxe commune à tout un corpus, c'est à dire le style musical ou le style de jeu d'un musicien. Ce genre d'analyse statistique peut être utilisé pour établir des regroupements de compositeurs et ainsi permettre à un programme après une écoute d'un morceau inconnu de déterminer son auteur. Mais aussi de générer de la musique *à la manière de* car l'action statistique permet d'effacer les gestes anecdotiques et de mettre en avant les gestes idiomatiques.

#### Exemple de modèle de Markov sur une mélodie simple

Nous allons ici construire une chaîne de Markov correspondant à la mélodie suivante, cela permettra de comprendre comment procéder pour de grands corpus.



On a donc quatre états : Do, Ré, Mi et Fa qui correspondent aux lettres de notre alphabet, on calcule maintenant le poids des transitions en calculons la probabilité de jouer un do après un ré, un ré avec un mi, et ainsi de suite :



A partir de cette chaîne de Markov on déduit une grammaire probabiliste simple qui nous dit par exemple que l'on ne peut pas jouer de Ré après un Fa. Si l'on procède sur de grands corpus, les gestes musicaux anecdotiques sont effacés et les gestes idiomatiques quant à eux sont accentués. On modélise donc la syntaxe du corpus. On peut ensuite utiliser ce modèle pour faire de la génération, il suffit de partir d'un état donné et de tirer une variable aléatoire  $X$  dans un ensemble  $E$  et de définir autant de sous ensemble de  $E$  que de transitions, si  $X \in E_i$  alors on choisit la transition  $i$ . On génère alors des compositions "à la manière de".

C'est cette méthode que Jon Gillick, Kevin Tang et Robert M. Keller des universités de Stanford, Cambridge et Harvey Mudd College ont utilisée.<sup>44</sup>

Leur but fut de générer des improvisations de solo de Jazz en utilisant des grammaires probabilistes (premiers travaux par Keller en 2007), l'article présente l'apprentissage automatique de telles grammaires. Ils partent de corpus de plusieurs musiciens de Jazz reconnus (Keith Jarett, Miles Davis, etc ...) et établissent des modèles probabilistes pour chacun d'entre eux afin de capturer des gestes idiomatiques de ces musiciens. En utilisant chacun de ces modèles pour la génération, le programme permet de générer des solos à la manière de célèbres musiciens et ce sur n'importe quelle grille d'accords. Un programme utilisant ce principe fut mis en place par Robert Keller de l'Université de Harvey Mudd College, il s'appelle Impro-Visor et on le trouve aisément sur internet.

Ces travaux peuvent être extrêmement intéressants d'un point de vue musicologique, ils permettent d'apprendre rapidement beaucoup de choses sur un style musical ou un jeu de musicien, de formaliser un langage de manière rapide et automatique mais aussi de générer des énoncés valides mais nouveaux, ce qui permet de tester le modèle mais aussi de procéder à des analyses sur ces énoncés afin d'en apprendre encore plus.

44. Gillick, Jon, Kevin Tang, and Robert M. Keller. « Machine learning of jazz grammars. » *Computer Music Journal* 34.3 (2010) : 56-66.

## 2 Méthode

### 2.1 Modélisation par contraintes

Recentrons nous à notre problème, le contrepoint étant formellement défini, nous voulons arriver à produire une équivalence à une grammaire générative le représentant afin de produire des énoncés valides selon son langage. La façon la plus simple pour cela est d'utiliser la modélisation par contraintes. Cela consiste à définir un ensemble de règles de telle manière à ce que l'on puisse ensuite produire des énoncés qui respectent toutes ces règles. Cette approche est dans sa finalité équivalente aux grammaires génératives car elle permet de produire tous les énoncés corrects d'un langage à partir de règles syntaxiques.

### 2.2 Construction du programme

Afin de pouvoir utiliser un modèle nous permette d'effectivement produire des énoncés musicaux, il nous faut créer ce modèle sur ordinateur. A cette fin nous allons coder un programme qui en tenant compte de toutes les règles de notre ouvrage permet de générer des partitions musicales d'une taille choisie. Ce programme a été réalisé en Java en utilisant la bibliothèque *javax.sound.midi* pour la gestion des fichiers midi.

Le programme initialise une partition abstraite contenant le cantus firmus sur lequel nous voulons travailler, ensuite il compose le contrepoint de proche en proche en partant du début de la partition. Il détermine à chaque étape toutes les notes valides et en choisie une au hasard. Pour cela toutes les règles du traité sont modélisées sous forme d'énoncés logique, et le programme ne garde que les notes qui permettent de satisfaire tous ces énoncés. Par exemple, la règle qui stipule que seul les tierces, sixtes, quintes et octaves sont autorisées sur les temps fors peut se représenter sous la forme du bloc d'instruction Si suivant :

```
1 Si (La note du cantus firmus est      distance d'une tierce ,
    quinte, sixte ou octave de la note que l'on teste) alors :
    continuer \\  
2 Sinon, la note n'est pas valide.
```

En java on peut considérer les notes comme des entiers positifs, le calcul d'intervalles se fait donc pas simple soustraction de deux entiers cette règles donne donc en Java le code suivant :

```
1  
2 if (note.intervalle(v.getNotes().elementAt(posHarmonie)) \% 12  
    != 0 && note.intervalle(v.getNotes().elementAt(posHarmonie  
    )) \% 12 !=7 && note.intervalle(v.getNotes().elementAt(  
    posHarmonie)) \% 12 !=3 && note.intervalle(v.getNotes().
```

```

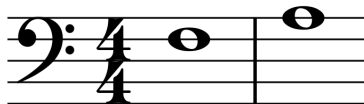
3     elementAt(posHarmonie)) \% 12 != 4 && note.intervalle(v.
4     getNotes().elementAt(posHarmonie)) \% 12 !=8 && note.
5     intervalle(v.getNotes().elementAt(posHarmonie)) \% 12 !=9){
        n.remove(note);
    }

```

## Maximisation

Ainsi il est possible de définir toutes les règles qui sont qualitatives, cependant il existe d'être de maximisations, les principales sont ; Il faut le plus possible utiliser le mouvement conjoint<sup>45</sup> et ; Il faut faire le plus de mouvements contraires<sup>46</sup> possible. Une solution simple aurait été d'écrire un mouvement conjoint ou contraire dès que cela est possible<sup>47</sup>, or cela ne nous garanti absolument pas leur maximisation, prouvons avec un contre exemple.

*Démonstration.* On souhaite écrire un contrepoint pour le cantus firmus suivant :



La paradigme glouton ne peut donner meilleur résultat que :



Or un optimum est :

---

45. Intervalle allant d'une note à une note voisine contiguë.

46. Lorsque deux parties se déplacent en sens inverse

47. Appelons ceci le paradigme glouton



Cela est du à la présence de nombreuses règles : par exemple il est obligatoire de faire un mouvement conjoint à la barre de mesure, or si l'on veut respecter la règles harmoniques (seul l'unisson, la tierce, la quinte, la quinte augmentée et la sixte sont autorisés sur les temps forts) le seul moyen pour le glouton qui a commencé par faire des mouvements conjoints est de rater un mouvement conjoint et rater le mouvement contraire. Alors que la méthode optimale est de rater le premier mouvement conjoint afin de maximiser plus tard les mouvements conjoints et contraires.

Ce paradigme n'est donc pas optimal dans notre cas.

□

De plus, ce paradigme est en pratique loin d'être optimal pour notre problème et crée des situations pathologiques récurrentes en diminuant considérablement l'ensemble des partitions possibles (il écrit un mouvement conjoint ou contraire dès qu'il le peut).

La solution idéale serait donc d'énumérer toutes les possibilités et de choisir l'optimum global. Or, le nombre de solutions est borné par la fonction :

$$f : \begin{cases} \mathbf{N} & \longrightarrow \mathbf{N} \\ n & \longmapsto 5^n \end{cases}$$

avec  $n$  la taille de partition en nombre de temps.

Il est donc impossible d'énumérer les contrepoints valides d'un cantus firmus de 20, 30 ou 40 temps. Il faut donc trouver une heuristique. L'idée est de trouver un maximum local en utilisant l'aléatoire : on fait tourner le programme sur un serveur de calcul pendant un temps indéfini proportionnel à la précision attendue, on paramètre le programme pour qu'il enregistre uniquement les partitions dont les paramètres à maximiser sont supérieurs à un seuil satisfaisant, on augmente ainsi considérablement la rapidité en diminuant les écritures fichier. Le programme classe toutes les partitions valides aux valeurs intéressantes dans une arborescence de fichiers, il ne nous reste plus qu'à choisir la meilleure partition au bout de l'intervalle de temps qui nous convient. Pour éviter les doublons la technique consiste à utiliser la fonction de hachage MD5 comme nom de fichier, ainsi on ne peut avoir deux fichiers contenant la même partition.



### 2.3 Test de la robustesse du modèle par test psychologique

Maintenant que nous pouvons générer des contrepoint théoriquement valides, il est intéressant de tester la validité de ce modèle, une première approche est de faire passer des tests à l'aveugle afin de voir si les personnes interrogées arrivent à reconnaître que les partitions sont écrites par un ordinateur.

Cet axe d'investigation nous permet de mieux comprendre les différences compositionnelles entre le programme et l'humain, puis d'évaluer le caractère universel de l'appréciation de ces différences. Pour ce faire il est nécessaire de procéder à une expérience sur un panel de personnes relativement grand. Et ensuite analyser ces résultats avec des outils statistiques efficaces<sup>48</sup>. Or, pour avoir des données intéressantes il faut réussir à mettre en place une expérience fiable et pertinente. Voici le protocole<sup>49</sup> :

---

48. Le traitement des données peut être fait avec le langage de programmation et environnement statistique R

49. Le test est disponible en annexe.

Numéro du test :

Nom :

Prénom :

Age :

Dans quel domaine travaillez-vous ?

Êtes-vous musicien ?      Oui ☐      Non ☐

Combien de temps passez vous à écouter de la musique par semaine?

$t < 3h$  ☐

$3h \leq t < 10h$  ☐

$t > 10h$  ☐







Connaissez vous le contrepoint ?      Oui ☐      Non ☐

### Partie I :

Cochez les adjectif qui vous semble le mieux correspondre au morceau.

Morceau 1	Morceau 2	Morceau 3
Cohérent <input type="checkbox"/>	Cohérent <input type="checkbox"/>	Cohérent <input type="checkbox"/>
Prévisible <input type="checkbox"/>	Prévisible <input type="checkbox"/>	Prévisible <input type="checkbox"/>
Audacieux <input type="checkbox"/>	Audacieux <input type="checkbox"/>	Audacieux <input type="checkbox"/>
Structuré <input type="checkbox"/>	Structuré <input type="checkbox"/>	Structuré <input type="checkbox"/>
Emouvant <input type="checkbox"/>	Emouvant <input type="checkbox"/>	Emouvant <input type="checkbox"/>
Froid <input type="checkbox"/>	Froid <input type="checkbox"/>	Froid <input type="checkbox"/>
Intelligible <input type="checkbox"/>	Intelligible <input type="checkbox"/>	Intelligible <input type="checkbox"/>

Cochez la couleur qui vous semble le mieux correspondre au morceau.

Morceau 1	Morceau 2	Morceau 3
 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>
 <input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>

### Partie II

Morceau 1	Morceau 2	Morceau 3
H <input type="checkbox"/> O <input type="checkbox"/>	H <input type="checkbox"/> O <input type="checkbox"/>	H <input type="checkbox"/> O <input type="checkbox"/>

L'expérience se déroule en deux parties.

### **Première partie de l'expérience**

Au début de l'expérience les candidats ne sont pas mis au courant de la problématique du travail. Il n'est en aucun cas mentionné la génération de partition par un ordinateur, seulement qu'ils devront écouter 3 variations d'un même thème musical. On procède à l'écoute de trois morceaux avec une des répartitions suivantes :

- { Deux morceaux écrits par le programme et un morceau écrit par un compositeur
- { Deux morceaux écrits par un compositeur et un morceau écrit par le programme

A chaque écoute le candidat est prié de choisir quelques adjectifs parmi une liste qui lui est proposée.

Ensuite il est demandé de choisir entre deux couleurs, celle correspondant le mieux à chaque morceau. Donc une et une seule réponse par morceau. Les couleurs sont alors choisies de telle manière à ce qu'elles aient peu de différences symboliques. L'idée est de permettre dans cette question de voir comme le candidat regroupe les trois morceaux : il y a deux modalités de couleurs et trois morceaux il faut donc qu'il mette la même couleur sur un moins un morceau.

### **Seconde partie de l'expérience**

On explique désormais au candidat le principe de l'expérience, la notion de génération aléatoire par contrainte à l'origine des partitions de programme. On lui demande alors de déterminer selon lui quels morceaux parmi ceux qu'il vient d'écouter sont générés par le programme et ceux écrits par un compositeur. Il peut s'il le désire écouter une deuxième fois chaque morceau.

Les deux parties sont réalisées avec le même échantillon de morceaux.

### **Motivation des choix**

Bien entendu les adjectifs ne sont pas choisis au hasard. Ils sont le fruit d'une réflexion prenant pour base l'écoute comparative approfondie des morceaux mais aussi et surtout de la composition fondamentale du programme. En effet, en ayant travaillé à la conception du programme j'ai une petite idée de ses procédés de compositions et de ses méthodes récurrentes. Il est donc intéressant pour moi de voir à quel degré mon échantillon de candidat pourra percevoir ces nuances. Il peut par exemple déterminer les différences subtiles et ainsi en déduire quels sont les morceaux écrits par l'ordinateur et ceux écrits par le compositeur. Mais il peut aussi sûrement saisir une différence dans le ressenti, dans les émotions transmises et dans l'esthétique de la composition sans avoir idée de quel morceau est composé par qui. De plus, la différence peut être perçue de manière extrêmement subtile, c'est l'intérêt de travailler avec le langage des couleurs. Le ressenti peut être tellement

abstrait que le candidat a du mal à mettre des mots sur cette différence, le recours aux couleurs peut alors être judicieux.

Grace à cette expérience on pourra avoir des pourcentages de discrimination entre l'origine humaine et computationnelle selon une échelle d'abstraction.

- Discrimination parfaite, le candidat sait reconnaître l'origine de la composition.
- Discrimination partielle, le candidat donne des adjectifs sensiblement différents selon l'origine de la composition.
- Discrimination abstraite, le candidat donne des couleurs différentes selon l'origine de la composition.

Il sera donc intéressant d'analyser et de comparer les résultats. On pourra aussi voir les adjectifs les plus souvent utilisés pour chaque modalité.

Il est évident que ces résultats seront à discuter en fonction du profil du candidat, des questions sont donc posées en début d'expérience : Le candidat est-il musicien, à quelle fréquence écoute-t-il de la musique, dans quel domaine travaille/étudie-t-il.

## 3 Analyse et synthèse

### 3.1 Résultats de l'expérience

#### 3.1.1 Les adjectifs dominants

On observe que les adjectifs dominants sont :

{ Structuré, Cohérent et Prévisible pour l'écriture de l'ordinateur.  
{ Structuré, Cohérent et Émouvant pour l'écriture du compositeur.

On observe donc que les candidats ont globalement trouvé tous les morceaux structurés et cohérents. Puis qu'ils ont trouvé les morceaux écrits par l'ordinateur plus prévisible que ceux écrits par le compositeur. Et enfin ils ont trouvé les morceaux écrits par le compositeur plus émouvants que les morceaux écrits par l'ordinateur.

On peut penser que cette tendance générale à trouver les morceaux structurés et cohérents vient directement de l'effet du contrepoint. En effet, ces principes de compositions étant soumis à des règles très strictes et prenant en compte un certain nombre de notes précédentes, la structure des morceaux s'en voit donc affectée. La cohérence quant à elle peut être expliquée par le fait que les règles ont un certain sens entre elle et façonne une esthétique générale.

Il est intéressant de remarquer cette différence de ressenti. En effet l'adjectif émouvant se rapporte directement à un sentiment humain. Cela veut dire que les candidats ont globalement trouvé les morceaux composés par le compositeur plus "humain" que les autres. Cependant comme on le verra tout à l'heure seulement 45% des participants ont réussi à reconnaître les morceaux de manière consciente (cocher la bonne case quand la question

leur est posé) cela veut dire qu'il émane des les morceaux composés par un humain un sentiment à priori plus humain que ceux composés par l'ordinateur, mais que les candidats ne font pas pour autant le rapprochement avec le fait que ces morceaux soient écrits par un humain et pas les autres. Quand au caractère prévisible des morceaux écrits par l'ordinateur, il est très étonnant. En effet l'ordinateur ne fait que respecter les règles de contrepoint, tout comme le compositeur. Or quand plusieurs choix s'offre à lui, l'ordinateur choisit une note au hasard. Il devrait donc théoriquement être complètement imprévisible par rapport au compositeur qui lui possède une culture musicale et peut faire des références<sup>50</sup> ou bien même connaît les enchainements qui fonctionnent. Cela peut peut-être s'expliquer par l'organisation des règles de contrepoint qui dans la majorité des cas (observé par l'expérience, car toutes les partitions sont générée aléatoirement et chaque candidat écoute des morceaux différents) donne une sensation "prévisible". Cela voudrait dire que le contrepoint impose des mécanisme compositionnels prévisibles. Cela est sûrement du au fait qu'un certains nombres de ces règles a été conservée tout au long de l'histoire de la musique. On ressent aujourd'hui même encore la présence de ces principes dans la musique populaire et aussi dans le jazz. Le fait d'explorer au hasard des chemins en suivant ces règles donne donc une impression de prévisible, alors que le compositeur contemporain doit s'efforcer à emprunter des chemins plus personnels pour composer un contrepoint plus intéressant et au ressenti moins prévisible.

### 3.1.2 Les degrés de discrimination

Dans notre échantillon on observe la répartition suivante :

- 45 % de Discrimination parfaite.
- 27 % de Discrimination partielle.
- 27 % de Discrimination abstraite.

On remarque donc que notre échantillon reconnaît relativement bien les morceaux composés par l'ordinateur et ceux composés par le compositeur. De plus, on remarque que le choix de l'utilisation des couleurs pour mettre en évidence la discrimination abstraite est en réalité un choix judicieux. En effet, en questionnant les candidats après l'expérience, j'ai remarqué que le choix de la répartition des couleurs se fait par comparaison entre les ressentis des morceaux et qu'un certains nombre de personnes rapprochent des morceaux (quasiment tout le temps par la modalité ordinateur/compositeur, sans même avoir été informés de cette dernière, les candidats sont alors persuadés que tous les morceaux sont composés par un compositeur). Cela nous informe sur le mécanisme de discrimination. Il existe donc une différence de ressenti générale entre nos deux modalités de composition. De plus le pourcentage de candidats n'ayant pas réussi à déterminer quels morceaux étaient composés par qui mais ayant réussi à regrouper les morceaux selon la modalité avec les

---

50. Voir le chapitre suivant sur l'analyse musicologique

couleurs n'est pas négligeable. On peut donc affirmer qu'il existe plusieurs niveau de reconnaissance de l'origine compositionnelle des morceaux, que l'on peut déterminer grâce à ce protocole. De plus on observe que ces différents degrés de reconnaissance sont décorrelés de tout facteur musical. On observe la répartition suivante :

Parmi les discriminations parfaite :

- 40 % écoutent moins de 3h de musique par semaine.
- 40 % écoutent entre 3 et 10h de musique par semaine.
- 20 % écoutent plus de 10h de musique par semaine.
- 0 % sont musiciens.

Parmi les discriminations abstraite :

- 100 % écoutent entre 3 et 10h de musique par semaine.
- 0 % sont musiciens.

Parmi les discriminations nulles :

- 33 % écoutent moins de 3h de musique par semaine.
- 33 % écoutent entre 3 et 10h de musique par semaine.
- 33 % écoutent plus de 10h de musique par semaine.
- 66 % sont musiciens.

On observe que les personnes écoutant globalement plus de musiques aurait moins de chances de reconnaître l'origine compositionnelle du morceau (ou plus abstraite). De plus on observe aussi que les musiciens ont plus de chance de faire fausse route : 66% pour les discrimination nulles contre 0% pour les discriminations totale et abstraite. Or étant donné la petite taille de notre échantillon, le petit nombre de musiciens dans l'échantillon (dont essentiellement des musiciens non professionnels) et les différences minimales dans les résultats. On ne peut conclure à une corrélation significative entre les degrés de reconnaissance et les facteurs musicaux. On conclue donc que les facteurs musicaux - écoute de musique, pratique d'un instrument - n'ont pas d'impact sur les degrés de reconnaissance de l'origine d'un morceau.

## **3.2 Analyse musicologique**

### **3.2.1 Analyse harmonique**

Il serait maintenant intéressant de se pencher sur les partitions et essayer de comprendre les mécanismes en oeuvre. Voici les partitions d'une expérience ayant été faite à un candidat.



FIGURE 1 – Première écoute  
Composé par l'ordinateur



FIGURE 2 – Seconde écoute  
Composé par l'ordinateur



FIGURE 3 – Troisième écoute  
Composé par un enseignant de contrepoint

On remarque tout d'abord une grande similarité entre ces énoncés, ce qui est dans un certain sens rassurant du fait qu'il sont composé à partir du même cantus firmus et qu'ils respectent tous le même ensemble de règles. On observe de manière indifférenciée l'utilisation du saut de la barre de mesure par des notes en valeur longues (ici de rondes) ainsi que l'utilisation des rythmes blanches, noires et croches. Les mélodies sont très similaires, en mouvement plutôt conjoints et contraires mais dans des registres différents mais toujours bien utilisés (pas d'ambitus trop grand ou de saut de registres). On remarque que tous les morceaux (de part la nature du cantus firmus) sont écrits en Do Ionien ce qui leur confère cette couleur modale inmanquable.

D'un point de vu harmonique le rencontre en début de mesure font entendre respectivement les intervalles suivants :

- 1 : Octave sous entendue, tierce, tierce, quinte, sixte, tierce retardée, tierce, tierce, sixte, octave
- 2 : Octave sous entendue, tierce, tierce, quinte, sixte, tierce retardée, tierce, quinte, sixte, octave
- 3 : Octave sous entendue, sixte, sixte retardée, sixte, tierce, sixte retardée, tierce, tierce, sixte retardée, octave

On remarque l'utilisation d'intervalles très pur, cependant la partition écrite par l'enseignant ne présentent aucune quinte, seulement des tierces, sixtes, directes ou retardées. Cela donne un coté plus systématique et assumé à l'écriture, contrairement aux partitions de l'ordinateur - qui de plus présentent quasiment les mêmes intervalles harmoniques pour des mélodies différentes - qui ont entendre tous les intervalles autorisés de manière équivalente. On ressent d'un un geste de compositeur dans les choix des intervalles harmoniques seulement dans la troisième partition. Choix qui serait de l'ordre de la sémantique, de l'intentionnalité du compositeur. Si l'on veut maintenant aller plus loin dans l'analyse on pourrait analyser ces petits textes avec des méthodes d'analyse tonale afin de voir ou non l'influence culturelle du musicien ; cette analyse n'a bien entendu pas d'autre sens ici, s'agissant de musique modale et non tonale. Nous retrouverions néanmoins les degrés de fonction suivants :

- 1 I VI V III II II V III VII I
- 2 I VI V III II VI V III VII I
- 3 I IV III I IV VI V III VII I

Notons l'utilisation dans tous les cas du cinquième degrés non fonctionnel et du troisième degrés (relatif mineur du cinquième) typique de la musique modale, n'engendrant pas d'attraction entre les accords. On retrouve néanmoins la même cadence VII I imposée par le mouvement de basse qui permet de terminer sur la tonique par un double mouvement



conjoint et contraire. On ne remarque pas à ce niveau de grande différences dans l'utilisation des degrés malgré un chemin harmonique toujours différent.

Une différence notable réside néanmoins dans forme générale de l'énoncé, la conduction des voix, les mélodies produites par l'ordinateur semble désordonnées, sans ligne directrices alors qu'elles sont pourtant parfaitement valides d'un point de vu mélodique et harmonique. Cela pose question sur la valeur culturelle de la musique, n'existerait-il pas des règles supplémentaire permettant de garantir un bon contrepoint en plus d'un contrepoint valide ? Ces règles serait-elle toujours de l'ordre de la syntaxe ? Voici la question à laquelle nous allons tenter de répondre pour conclure ce document.

Du fait que nous avons réussi à produire des énoncés musicaux qui respectent toutes les règles du contrepoint mais qui quand on les analyse à l'échelle de la forme ou par des méthodes psychologiques ne répondent pas tout à fait à ce que l'on pourrait attendre de lui. Il se trouve qu'il semble difficile d'ajouter localement des règles pour produire des énoncés plus crédibles. Le fait par exemple d'utiliser que des intervalles de tierces et sixtes ne peut émaner que de la pensée intentionnelle d'un musicien ; on pourrait intégrer ce procédé dans le programme mais il serait alors coder en dur et correspondrait à l'intentionnalité du programmeur et ne ferait que simuler des états mentaux d'intentionnalité. On est donc dans le domaine de la sémantique, et non plus de la syntaxe. Il apparait donc qu'un genre musical se définit par une syntaxe et par une sémantique (comme nous l'avons vu au préalable) et que comme dans le langage naturel la syntaxe ne suffit par à produire des énoncés bien formés : ils doivent être bien formés car régis par une syntaxe mais doivent aussi avoir un sens dans un contexte sémantique bien précis. Nous allons donc tenter de montrer ici que les notions de de syntaxe et sémantique jusqu'ici séparées en deux objets bien distinct tendrait à se rassembler en une notion bien complexe. Nous allons pour expliciter notre pensée nous appuyer sur un exemple présents dans les partitions étudiées. C'est ce que nous appellerons la *divergence de registres*.

### 3.2.2 Le registre au sens littéraire

On appelle registre littéraire, tonalité ou ton, l'ensemble des caractéristiques d'un texte qui provoquent des effets particuliers (émotionnels ou intellectuels) sur le spectateur ou le lecteur. On peut dans un même texte rencontrer plusieurs registres et aucun registre n'est lié exclusivement à un genre littéraire.

La remarque la plus importante que l'on peut faire à l'écoute de nos textes musicaux se rapporte à cette notion de registre. En effet, n'importe quel morceau de musique peut se rattacher à un registre -la notion est peut être plus générale qu'au sens littéraire - pour évoquer ces registres le compositeur utilise des procédés particuliers. Nous nommerons par la suite *fonction sémantique* le mécanisme reliant le procédé à l'effet provoqué. Or un programme ne connaît en aucun cas la fonction sémantique des procédés qu'il peut utiliser, il peut donc tout à fait utiliser par pur hasard un procédé couramment utilisé pour provoquer un effet précis. Cependant, l'utilisation de ce procédé n'étant pas contrôlée, le

programme peut donc utiliser très vite un autre procédé provoquant un effet totalement différent. De temps en temps ce mécanisme peut entraîner des changements de registres inattendu, paradoxaux. Appelons cette notion la *divergence de registre*. Cette divergence de registre est omniprésente dans la musique informatique composée par contraintes avec des règles destinées à la pédagogie musicale, en effet cette notion de divergence de registre est trop intuitive pour un compositeur pour avoir à être formalisée. Cependant, il est intéressant de réfléchir à l'origine de la fonction sémantique d'un procédé. Comment se fait-il que la modélisation de cette notion soit si malaisée ? Quels sont les facteurs qui affectent cette notion ?

### 3.2.3 Evolution de la *fonction sémantique* et concept de *connotation*

Roland Barthes nous donne dans son explication d'une publicité panzanni un exemple de relation entre *connotation culturelle* et fonction sémantique. En effet, c'est la connotation méditerranéenne des produits frais présents dans le panier qui donne du sens à leur utilisation dans la photographie<sup>51</sup>. Et si le mécanisme principal de *l'évolution sémantique* des procédés du langage musical était le concept de connotation ? Par exemple, serait-ce la connotation d'un procédé qui lui permettrait d'évoquer la tristesse ? Pour pouvoir répondre essayons de réfléchir au mécanismes évolutifs de la connotation d'un procédé. Imaginons un final extrêmement tragique d'un opéra du XIXème siècle. Imaginons maintenant que le compositeur essaie lors de ce final un procédé alors méconnu, tout les éléments vont dans le sens de la tragédie, excepté le nouveau procédé. Le final n'en sera pas moins tragique, cependant ce procédé alors utilisé sera surement coloré de cette dimension tragique. S'il est réutilisé isolé plus tard il aura donc une fonction sémantique qui pourra tendre vers le tragique. Prenons un autre exemple, plus pragmatique cette fois. Les procédés couramment utilisés dans le genre cinématographique d'horreur avaient-ils cette connotation d'horreur avant leur emploi dans ce genre de films. Où bien est-ce la surabondance de leur utilisation qui leur confère cet effet ? Il semble que l'utilisation de la note pédale au violon avait à l'époque romantique une toute autre signification que celle que les bandes originales de films d'horreur lui donne. Son utilisation dans ce genre cinématographique est essentiellement celui de l'attente angoissée et interminable avant l'événement tragique. On retrouve ce procédé dans *Le locataire* de Roman Polanski en 1976, notamment dans la séquence *le délire hallucinatoire de Trelkovsky*, il se retrouve aujourd'hui systématiquement dans les film d'horreurs à gros budget et même sous des formes exagérées comme par exemple dans *La habitación del niño* de Álex de la Iglesia.

La sémantique en musique serait donc le fait de la culture, de l'habitude, de la connotation des agencements de signes. Les procédés dans l'histoire des pratiques musicales se colorent et nous permettent de leur donner du sens. Nous voyons ainsi grâce à l'analyse des partitions générées par notre modèle génératif que syntaxe est sémantique forment finale

---

51. Voir : *Une publicité pour les pâtes Panzani. Ce que l'on doit comprendre se limite-il à ce que l'on voit ?* Aisément trouvable sur internet.

un seul et même objet, un objet de culture. En effet en musique il ne peut y avoir de sémantique sans syntaxe car la sémantique ne peut par essence exister sans elle. Mais, il ne peut non plus y avoir syntaxe sans sémantique car comme nous l'avons vu plus haut la modélisation syntaxique d'un langage ne permet pas pour autant que créer des énoncés valides car non cohérent.

## Conclusion

Nous avons dans ce document vu que la musique peut se poser en langage pour le musicologue de part sa capacité à fédérer les hommes et transmettre des idées esthétiques. De plus, on peut d'une certaine façon par des exemples précis trouver un moyen de trouver une équivalence aux notions de syntaxe et sémantiques en musique. Il nous ait donc paru possible de pouvoir utiliser des outils issus de la linguistiques pour travailler sur le langage musical. C'est le cas des grammaires de Chomsky que vous avons défini et qui nous ont permis de construire un modèle génératif du contrepoint rigoureux. Cette démarche nous a permis de mieux comprendre le contrepoint dans son aspect formel mais aussi d'étudier les partitions qu'il était capable de générer. Puis de grâce aux distinctions entre sémantiques que nous avons faites, notamment grâce à l'expérience de pensée de John Searle, nous pouvons juger, après analyse des partitions, de l'importance de la sémantique dans la structure même du langage. Ce qui nous paraissait alors être de l'ordre de la syntaxe - la forme et la structure - se trouve maintenant du côté sémantique, et nous permet de repenser cette distinction ainsi que les mécanismes évolutifs la régissant. Si bien qu'après cette étude nous pouvons mieux comprendre comme la musique s'érige en langage et les interactions que peuvent entretenir syntaxe et sémantique. Ce travail, avant tout théorique et méthodologique, permet d'ouvrir la porte à une nouvelle façon de penser la sémiologie musicale, par l'utilisation de modèles génératifs. Nous avons pu ici mettre en évidence la notion de divergence de registres dans le contrepoint, mais poussée plus cette méthode pourrait nous permettre d'en apprendre plus sur les langues musicales, voire même sur le langage musical.

## Bibliographie

- Claude Abromont, *Guide de la théorie de la musique*, Fayard, 2001, 609 p.
- André Boucourechliev, *Le langage musical*, Fayard, Les chemins de la musique, 1993, 186 p.
- Noël Gallon et Marcel Bitsch, *Traité de contrepoint*, Durand et Cie, 1964, 130 p.
- Olivier Carton, *Langages formels, calculabilité et complexité*, Vuibert, 28 octobre 2008, 236 p.
- Noam Chomsky, *Three models for the description of language*, IRE Transactions on Information Theory ( Volume : 2, Issue : 3), Septembre 1956.
- Catherine FUCHS, « LINGUISTIQUE - Théories ». In Universalis éducation [en ligne]. Encyclopædia Universalis, consulté le 29 mai 2017. Disponible sur <http://www.universalis-edu.com/encyclopedie/linguistique-theories/>
- Gillick, Jon, Kevin Tang, et Robert M. Keller. *Machine learning of jazz grammars*, Computer Music Journal 34.3, 2010, pp. 56-66
- Daniel Jurafsky et James H. Martin, *Speech and Language Processing.*, Chapitre 11 : Formal Grammars of English, 30 janvier 2017.
- Jacqueline Léon. *Les chaînes de Markov. Parcours d'un 'modèle' au fondement de la mathématisation et de l'automatisation de la linguistique.*, Modélisations et sciences humaines. Figurer, interpréter, simuler, L'Harmattan, pp.121-146
- Jean-Jacques Nattiez, *Fondements d'une sémiologie de la musique*, Union Générale d'Édition, 1975, 448 p.
- Edward Sapir, *Language, a introduction to the Study of Speech*, Harcourt, A harvest book, Brace and World, Inc., New York, 1949.
- Ferdinand de Saussure, *Cours de linguistique générale*, Payot, Paris, 1965.

Searle J.R. *Minds, brains and programs*. The Behavioral and Brain Sciences, 3, 1980.

Jean-Pierre Zarader, *Dictionnaire de philosophie*, Ellipses poche, 2014, 805 p.

## Annexes

### Pertinence du modèle par contraintes

Je vais dans cette partie tenter de montrer que le langage contrapuntique se prête bien à l'utilisation de la modélisation par contraintes de part sa nature de langage rationnel (qui prouve sa calculabilité d'après la théorie des langages formels). Pour cela, montrons que l'on peut construire un automate fini déterministe le reconnaissant.

*Démonstration.*

- i ) On définit l'alphabet  $\Sigma$  comme le produit cartésien des différents mouvements permis (On les notera par des chiffres romains : V est la quinte ascendante, -III est la tierce descendante, etc ...) avec l'ensemble des différentes durées de notes possibles. Tout contrepoint étant transposable, on peut se restreindre à la gamme de Do majeur et n'autoriser que les mouvements entiers, le nombre de mouvements permis est défini par l'ambitus maximal<sup>52</sup>.
- ii ) Toutes les règles sont présentées sous formes d'interdictions ou d'obligations de certaines configurations -à l'exception de deux règles qui sont des maximisations que l'on traitera plus tard-, par exemple :  $\mathbb{I}$  On ne peut écrire plusieurs quintes consécutives  $\mathbb{I}\mathbb{I}$  ou  $\mathbb{I}$  une quinte ascendante s'accompagne d'un mouvement descendant<sup>53</sup>  $\mathbb{I}\mathbb{I}$ . Il est clair que ces configurations peuvent se représenter par un nombre fini d'états d'un automate étant donné que le nombre de règles est fini. Chaque état possède un état de sortie qui correspond soit au mouvement obligatoire, soit à  $\Sigma \setminus \{\text{mouvements interdits}\}$

Si l'on ne veut pas forcément un automate minimal, il est aisé de construire un automate tel que chaque état est une configuration possible (mouvement permis), on y ajoute quelques états pour définir les règles qui prennent en compte les  $n$ -derniers mouvements.

---

52. L'ambitus en musique est l'écart entre le note la plus haute et la plus basse.

53. Cette règle n'est pas exprimée de la sorte, je l'utilise pour l'exemple.

iii ) Tous les états sont acceptants, car un contrepoint peut faire la taille que l'on désire, il peut donc terminer dans n'importe quel état.

L'automate entier ne peut être représenté ici, cependant voici un exemple avec les règles précédentes, en omettant la durée des notes.

□

L'appartenance d'un morceau au langage contrapuntique est donc décidable. On a ainsi montré qu'il est théoriquement possible de modéliser formellement les règles du contrepoint, qu'en est-il en pratique ?