



ANR-14-CE24-0002-01

## Projet DYCI2, WP2 Apprentissage interactif de structures musicales, WP2.3 Apprentissage de structures multi-échelles

### Rapport de livrable :

### L2.3 Apprentissage de structures multi-échelles.

Livrable	Date	Contributeurs	Rédacteurs	Contenu
L2.3 Version 01	Février 2018	K. Déguernel, E. Vincent (Inria), G. Assayag (Ircam-STMS)	K. Déguernel	Maquette Logicielle, Rapport scientifique

### Résumé

*Ce document décrit trois contributions effectuées dans la thèse de Ken Déguernel : a) un cadre de modélisation de la structure temporelle hiérarchique d'une improvisation par une grammaire formelle ; b) une méthode d'apprentissage automatique d'une telle grammaire basée sur des méthodes probabilistes de sélection de séquences de mots partageant une information mutuelle ; c) une méthode permettant de générer des improvisations temporellement cohérentes à partir d'une telle grammaire. Ces méthodes ont été évaluées pour des scénarios d'analyse et de génération par des sessions d'écoute avec des musiciens de jazz professionnels.*

### Adresse du livrable logiciel

DYCI2\_WP2\_L2.3.zip

sur

<https://forge.ircam.fr/p/Dyci2/>

# Livrable L2.3

Ce rapport correspond au chapitre 6 de la thèse de Ken Déguernel [Déguernel, 2018] et à l'article de conférence *Sound and Music Computing 2017* "Generating Equivalent Chord Progressions to Enrich Guided Improvisation : Application to Rhythm Changes" [Déguernel et al., 2017].

Nous présentons dans ce livrable les travaux effectués relatif à l'organisation à plusieurs niveaux temporels d'une improvisation. Par exemple, dans le cas d'un morceau de jazz, la grille d'accords est souvent structurée à plusieurs échelles temporelles différentes. Certains groupes d'accords (petite échelle) peuvent former des fonctions tonales ou modales (moyenne échelle) et ces fonctions peuvent être organisées en différentes sections (grande échelle). Nous nous intéressons ici au cas de l'improvisation guidée par un guidage structurel avec l'utilisation d'un scénario temporel [Nika et al., 2017]. Notre objectif est premièrement de pouvoir représenter l'organisation temporelle hiérarchique d'un scénario par un *scénario multi-niveaux*<sup>1</sup>. Pour cela, nous proposons d'utiliser une grammaire syntagmatique [Chomsky, 1972] permettant de générer plusieurs variations d'une même structure globale et de représenter l'organisation hiérarchique à plusieurs niveaux temporels. Nous voulons ensuite utiliser cette information multi-niveaux lors du processus de génération de l'improvisation à l'aide de nouvelles heuristiques adaptées à cette information. Ces travaux se veulent comme une extension des travaux proposés dans *ImproteK* par Nika [2016]; Nika et al. [2017].

Nous présentons dans la partie 1 comment une grammaire basée sur une analyse linguistique en constituants peut représenter la structure hiérarchique d'un scénario temporel pour constituer des scénarios multi-niveaux. Nous construisons, puis évaluons, une telle grammaire avec des musiciens professionnels pour un cas particulier de scénario multi-niveaux : une grille d'accords de jazz appelée anatole. Ensuite, dans la partie 2, nous proposons une méthode d'apprentissage automatique d'une telle grammaire basée sur des méthodes probabilistes de sélection de séquences de mots partageant une information mutuelle [Zitouni et al., 2000]. Dans la partie 3, nous proposons une nouvelle heuristique pour générer une improvisation sur un scénario multi-niveaux. Cette heuristique est basée sur les principes d'anticipation de scénario proposés par Nika et al. [2017]. Finalement, dans la partie 4, nous présentons une évaluation de notre méthode de génération effectuée lors de sessions d'écoute avec des musiciens de jazz professionnels.

---

1. Par la suite, nous choisissons le terme « niveau » plutôt qu'« échelle » pour décrire l'organisation temporelle afin d'éviter la confusion avec la notion d'échelle de notes.

## 1 Utiliser une grammaire pour modéliser une structure multi-niveaux

Dans cette partie, nous présentons comment utiliser une grammaire pour représenter l'organisation hiérarchique d'une grille d'accords en plusieurs niveaux temporels. Dans la partie 1.1, nous présentons le type de grammaire que nous allons utiliser : les grammaires syntagmatiques. Puis, dans la partie 1.2, nous construisons, avec un musicien professionnel, une grammaire syntagmatique pour représenter l'aspect multi-niveau d'une grille de jazz très populaire : l'anatole. Cette grammaire est ensuite évaluée par des musiciens professionnels dans la partie 1.3.

### 1.1 Grammaire syntagmatique

Les définitions présentées dans cette partie sont adaptées de Chomsky [1956, 1972] et de Hopcroft & Ullman [1979].

Soit un ensemble de symboles  $X$ , on note  $X^*$  l'ensemble des séquences finies d'éléments de  $X$ . Une grammaire  $G = (N, \Sigma, R, s)$  est définie par :

- un ensemble  $N$  de symboles non terminaux,
- un ensemble  $\Sigma$  de symboles terminaux, disjoint de  $N$ ,
- un élément particulier  $s \in N$  appelé axiome,
- un ensemble fini de règles de réécriture

$$R \subset (N \cup \Sigma)^* N (N \cup \Sigma)^* \rightarrow (N \cup \Sigma)^* . \quad (1)$$

Une règle de réécriture, généralement notée  $u \rightarrow v$  peut être interprétée comme une instruction signifiant « remplacer  $u$  par  $v$  ». Le langage  $L(G)$  généré par une grammaire  $G$  est l'ensemble des séquences de symboles terminaux pouvant être créé à partir de l'axiome  $s$  en utilisant les règles de réécriture de  $R$  :

$$L(G) = \{w \in \Sigma^* \mid s \xrightarrow{*} w\} , \quad (2)$$

où  $\xrightarrow{*}$  représente l'utilisation d'un ensemble fini de règles de réécriture de  $R$  pour réécrire  $s$  en  $w$ .

Une *grammaire syntagmatique* est un type particulier de grammaire, présentée dans Chomsky [1956]. Ces grammaires se basent sur une description linguistique d'un langage au niveau syntaxique par une analyse en constituants, c'est-à-dire par une décomposition des fonctions linguistiques au sein d'une structure hiérarchique. Chomsky [1972] propose un exemple de grammaire syntagmatique pour la construction de phrases simples en anglais, présenté ici dans la Grammaire 1.

---

**Grammaire 1** Exemple de grammaire syntagmatique pour la construction de phrases en anglais

---

- 1:  $Phrase \rightarrow GN + GV$
  - 2:  $GN \rightarrow Article + Nom$
  - 3:  $GV \rightarrow Verbe + GN$
  - 4:  $Article \rightarrow a, the...$
  - 5:  $Nom \rightarrow man, ball...$
  - 6:  $Verbe \rightarrow hit, took...$
-

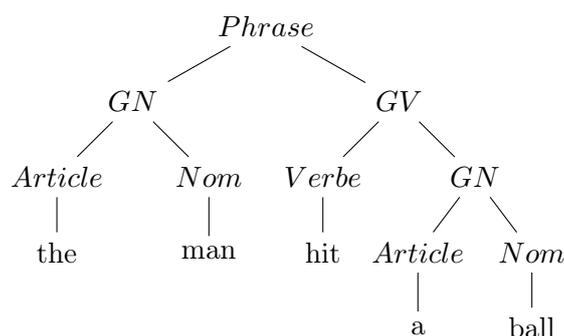


FIGURE 1 – Diagramme de la dérivation de la phrase « the man hit a ball ».

Dans cet exemple, les symboles non terminaux sont écrits en italique et les symboles terminaux en police normale. La règle 1 peut être lue comme « réécrire *Phrase* par *GN* + *GV* », c'est-à-dire qu'une phrase est constituée d'un *groupe nominal* suivi par un *groupe verbal*. Une interprétation similaire peut être effectuée pour les règles 2 et 3. La règle 4 spécifie que le symbole non terminal *Article* peut être réécrit par l'un des symboles terminaux proposés. Une interprétation similaire peut être effectuée pour les règles 5 et 6.

Une *dérivation* est la séquence de règles de réécriture utilisée pour obtenir une certaine phrase. Par exemple, pour obtenir la phrase « the man hit a ball », la dérivation présentée dans le Tableau 1 a été utilisée. Les numéros à droite de chaque ligne indiquent la règle de réécriture utilisée pour construire cette ligne à partir de la ligne précédente.

<i>Phrase</i>	
<i>GN</i> + <i>GV</i>	1
<i>Article</i> + <i>Nom</i> + <i>GV</i>	2
<i>Article</i> + <i>Nom</i> + <i>Verbe</i> + <i>GN</i>	3
the + <i>Nom</i> + <i>Verbe</i> + <i>GN</i>	4
the + man + <i>Verbe</i> + <i>GN</i>	5
the + man + hit + <i>GN</i>	6
the + man + hit + <i>Article</i> + <i>Nom</i>	2
the + man + hit + a + <i>Nom</i>	4
the + man + hit + a + ball	5

TABEAU 1 – Dérivation pour obtenir la phrase « the man hit a ball » à partir de la Grammaire 1.

Une dérivation peut être également représentée sous forme de diagramme. La Figure 1 montre le diagramme pour la dérivation utilisée pour la phrase « the man hit a ball ». Ce diagramme est moins riche en informations que la dérivation, car l'ordre d'application des règles de réécriture n'apparaît pas. Le diagramme ne retient de la dérivation que ce qui est essentiel pour déterminer la structure syntagmatique. Il permet alors de voir clairement la structure syntaxique hiérarchique de la phrase et une visualisation de l'analyse en constituants.

A	I VI-   II- V <sup>7</sup>   I VI-   II- V <sup>7</sup>
A	I I <sup>7</sup>   IV IV-   I VI-   II- V <sup>7</sup>
A	I VI-   II- V <sup>7</sup>   I VI-   II- V <sup>7</sup>
A	I I <sup>7</sup>   IV IV-   I V <sup>7</sup>   I
B	III <sup>7</sup>   III <sup>7</sup>   VI <sup>7</sup>   VI <sup>7</sup>
B	II <sup>7</sup>   II <sup>7</sup>   V <sup>7</sup>   V <sup>7</sup>
A	I VI-   II- V <sup>7</sup>   I VI-   II- V <sup>7</sup>
A	I I <sup>7</sup>   IV IV-   I V <sup>7</sup>   I

FIGURE 2 – Grille d’accords du thème *I Got Rhythm* de George Gershwin. Les accords sont marqué en degrés par rapport à la tonalité initiale du morceau.

## 1.2 Une grammaire syntagmatique pour les anatoles

Afin de montrer l’utilisation d’une grammaire syntagmatique pour créer un scénario multi-niveaux, nous avons décidé de créer une telle grammaire pour une progression d’accords très utilisée dans le jazz traditionnel et le *bebop* : l’*anatole*. Nous décrivons en premier lieu la structure d’un anatole du point de vue de l’analyse musicale (une description plus précise peut-être trouvée dans [Siron, 2015]), puis nous formalisons une grammaire syntagmatique appropriée représentant cette analyse hiérarchique de la progression d’accords.

### Analyse de l’anatole

L’anatole (en anglais *rhythm changes*) est une progression d’accords de 32 mesures, basée sur le thème *I Got Rhythm* de George Gershwin. Le terme *rhythm changes* est en fait un diminutif de « *Chord changes of I Got Rhythm* » (en français « progression d’accords de *I Got Rhythm* »). Le terme français « anatole » a été inventé par Jean-Claude Fohrenbach. La Figure 2 montre la progression d’accord originale de *I Got Rhythm*.

La structure globale de l’anatole consiste en une forme *AABA*, où la section *B* appelée le pont rentre en contraste avec les sections *A*.

- D’un côté, les sections *A* sont des sections de huit mesures avec des accords changeant rapidement avec généralement deux accords par mesure et restant proches de la tonalité initiale. Ces sections sont constituées de :
  - une série de deux *cellules-anatoles* sur les mesures 1&2 puis 3&4. La cellule-anatole est une fonction tonale de deux mesures basée sur un fragment du cycle des quintes. Sa forme la plus caractéristique est I VI- II- V, mais elle est régulièrement modifiée par les substitutions classiques (par exemple, I I II-V<sup>7</sup>, ou III- VI<sup>7</sup> II- V<sup>7</sup>, etc.).
  - une *cellule-christophe* sur les mesures 5&6 faisant sonner le IVème degré (sous-dominante). Le terme « christophe » a également été inventé par Jean-Claude Fohrenbach d’après le thème *Christopher Columbus*. Sa forme caractéristique est I I<sup>7</sup> IV IV-, mais elle est également régulièrement modifiée par les substitutions classiques. Par exemple, V- I<sup>7</sup> IV ♯IVo, ou I I<sup>7</sup> IV<sup>7</sup> ♭VII<sup>7</sup>, etc.

- une dernière cellule-anatole sur les mesures 7&8. À l'exception du premier  $A$ , cette cellule-anatole peut être remplacée par une *cadence-boucle* ayant pour but soit de conclure dans l'accord de tonique soit d'anticiper la fonction d'accord suivante. Une cellule-anatole est alors un cas particulier de cadence-boucle. Par exemple, une cadence-boucle peut être  $\text{II- } \mathbf{V}^7 \text{ I I}$ .
- D'un autre côté, la section  $B$  est une section de huit mesures avec une progression plus lente où chaque accord est généralement joué sur deux mesures, basé sur des accords de dominante suivant le cycle des quintes ( $\text{III}^7 \text{ VI}^7 \text{ II}^7 \text{ V}^7$ ) donnant une sensation de changement tonal. Les improvisateurs ont tendance souligner ces changements en insistant sur les notes guides (c'est-à-dire la tierce et la septième) de ces accords de dominante. Encore une fois, des substitutions peuvent être effectuées, par exemple, les deux mesures de  $\mathbf{V}^7$  peuvent être remplacées par une mesure de  $\text{II-}$  suivie d'une mesure de  $\mathbf{V}^7$  (ou de  $\text{bIII}^7$ ).

L'anatole est un cas d'étude intéressant pour notre application, car il existe une grande quantité de variations autour de cette progression d'accords. C'est d'ailleurs l'attrait principal pour les musiciens qui peuvent modifier la progression à la volée et effectuer différentes substitutions au cours de l'improvisation, tant que la forme globale et les différentes fonctions sont présentes. La progression d'accords jouée peut alors être différente à chaque itération de la grille. Parmi les anatoles les plus célèbres, on peut citer *Anthropology*, *Dexterity*, *Oleo*, *Rhythm-A-Ning*, *The Eternal Triangle*...

Utiliser une grammaire syntagmatique pour analyser et générer des anatoles semble alors approprié. En considérant les accords, les fonctions tonales et les sections comme constituants, nous pouvons créer une grammaire syntagmatique représentant la structure hiérarchique des anatoles et où les accords sont les symboles terminaux.

## Construction de la grammaire

Afin de créer une grammaire syntagmatique pour l'anatole, nous avons analysé avec Pascal Mabit, musicien professionnel et professeur de jazz, l'ensemble des anatoles issus du corpus de l'*Omnibook* [Déguernel et al., 2016]. Ce sous-corpus contient 26 variations différentes d'anatoles.

Nous définissons d'abord les notations utilisées pour les différents constituants décrits dans la partie précédente et jugés nécessaires pour la création de la grammaire :

- On note  $\tau$  une cellule-anatole. On note également  $\tau_{\text{I}}$  le sous-ensemble des cellules-anatoles commençant par le degré I ( $\tau_{\text{I}} \subset \tau$ ).
- On note  $\sigma$  une cellule-christophe.
- On note  $\omega$  une cadence-boucle (à noter que  $\tau \subset \omega$ ).
- On note  $\delta_{\text{X}}$  les fonctions d'accords de dominante sur le degré X trouvées dans la section  $B$ .

La grammaire syntagmatique pour les anatoles que nous avons créée à partir de l'analyse du sous-corpus de l'*Omnibook* est présentée dans la Grammaire 2.

- La règle 1 décrit la structure en  $AABA$  de l'anatole. Ces sections de huit mesures sont les plus grands constituants après la progression d'accords complète.
- Les règles 2 et 3 décrivent la composition d'une section  $A$  en quatre fonctions tonales de deux mesures. Une section  $A$  commence toujours avec un élément de  $\tau_{\text{I}}$  pour affirmer la tonalité au début de la section avec un accord de degré I. Ensuite, un autre  $\tau$  est joué, suivi par le passage plagal  $\sigma$ . La différence entre un  $A_1$  et un  $A_2$  se situe sur les deux dernières mesures. D'un côté,  $A_1$  se termine par une cellule-anatole

---

**Grammaire 2** Grammaire syntagmatique pour les anatoles

---

- 1:  $Anatole \rightarrow A_1 + A_2 + B + A$
  - 2:  $A_1 \rightarrow \tau_I + \tau + \sigma + \tau$
  - 3:  $A_2 \rightarrow \tau_I + \tau + \sigma + \omega$
  - 4:  $A \rightarrow A_1, A_2$
  - 5:  $B \rightarrow \delta_{III} + \delta_{VI} + \delta_{II} + \delta_V$
  - 6:  $\tau_I, \tau, \sigma, \omega, \delta_{III}, \delta_{VI}, \delta_{II}, \delta_V$  sont appris sur le corpus.
- 

$\tau$ , de l'autre côté  $A_2$  se termine par une cadence-boucle  $\omega$ .

- La règle 4 indique que la dernière section  $A$  peut soit être un  $A_1$  ou un  $A_2$ .
- La règle 5 décrit la composition d'une section  $B$  en quatre fonctions tonales de deux mesures. Chaque  $\delta$  décrit un déplacement tonal sur le cycle des quintes en suivant les degrés III VI II et V.
- La règle 6 spécifie que le passage des symboles non terminaux (ici fonctions tonales) aux symboles terminaux (ici accords) est appris sur un corpus d'apprentissage. Chaque fonction tonale est composée d'une séquence de quatre accords d'une durée de deux temps (avec la possibilité de répétition). Apprendre les symboles terminaux sur un corpus permet de considérer un ensemble large de possibilités pour chaque fonction et d'effectuer une forme de modélisation du style du corpus pour la génération de scénario multi-niveaux.

La Figure 3 montre le diagramme d'une dérivation de cette grammaire pour l'un des anatoles de l'*Omnibook* sur le thème *Celerity*.

### 1.3 Évaluation de la grammaire

Afin d'évaluer la Grammaire 2 pour la représentation des anatoles, nous avons généré avec elle 30 dérivations d'anatoles. Les progressions d'accords multi-niveaux ont été analysées par deux musiciens de jazz professionnels :

- Pascal Mabit, avec qui cette grammaire a été construite,
- Louis Bourhis, contrebassiste de jazz qui n'était pas impliqué dans le processus et qui par conséquent a analysé les progressions d'accords d'un point de vue purement musicologique.

La Figure 4 montre un exemple d'anatole généré par la grammaire syntagmatique.

Tous les anatoles générés avec la grammaire syntagmatique ont été validés par les deux musiciens comme des anatoles valides. Cependant, les progressions d'accords générées doivent être vues comme des réalisations d'anatoles effectuées en pratique lors de l'accompagnement d'une improvisation. Ceci est cohérent avec le corpus d'apprentissage utilisé dans le sens où les progressions d'accords notées dans l'*Omnibook* essaient de se rapprocher à chaque grille à ce qui a été joué en accompagnement des solos de Charlie Parker. Pour la composition automatique de grilles d'anatoles pour un thème de jazz, il serait judicieux d'utiliser des contraintes de parallélisme [Giraud et al., 2014] pour assurer des symétries entre les différents  $A$  d'une même grille.

De plus, les anatoles générés avec cette grammaire couvrent l'ensemble des possibilités des anatoles traditionnels de type *bebop*, ce qui correspond au corpus d'apprentissage utilisé. Aucune variation importante d'anatole de ce style n'a été jugée manquante, validant la capacité de modélisation du style de l'apprentissage des symboles terminaux de la grammaire. Cependant, cette grammaire ne permet pas de générer des formes plus modernes

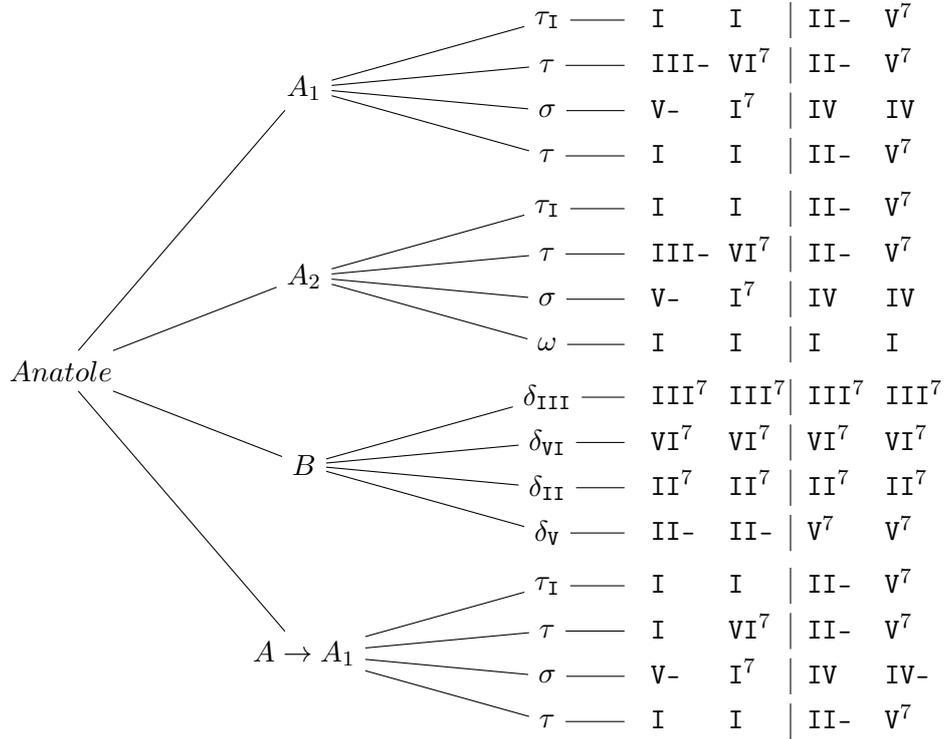


FIGURE 3 – Diagramme de la dérivation d’anatole du thème *Celerity* de Charlie Parker (issu de l’*Omnibook*. Chaque accord dure deux temps.  $A \rightarrow A_1$  représente l’application de la règle 4 de la grammaire 2.

d’anatoles comme par exemple celle de *The Eternal Triangle* par Sonny Stitt qui possède une structure différente pour la section  $B$ , ou la version de *The Theme* par Lee Morgan dans l’album *Straight Ahead* qui possède des substitutions d’accords *post-bebop*. La figure 5 montre la section  $B$  de *The Eternal Triangle*. Cette section conserve l’aspect contrastant du pont avec la section  $A$  et la notion de déplacement tonaux sur le cycle des quintes, mais sa structure diffère du pont de l’anatole traditionnel et utilise beaucoup de substitutions tritoniques. Il serait intéressant d’étendre la grammaire pour les anatoles pour des progressions d’accords de styles plus variés.

Ayant montré qu’une grammaire syntagmatique était une représentation efficace de l’organisation structurelle hiérarchique d’un scénario, nous souhaitons effectuer un apprentissage automatique de cette structure.

## 2 Apprentissage de structures multi-niveaux

Dans cette partie, nous proposons une méthode pour effectuer un apprentissage automatique de structures multi-niveaux à partir d’un corpus d’apprentissage (dans notre cas des séquences d’accords). Pour cela, nous souhaitons utiliser des méthodes d’induction de grammaire [de Marcken, 2015; Pate & Johnson, 2016]. Notre idée se base sur les travaux de Zitouni et al. [2000] sur la sélection de séquences de mots pour effectuer une compression de l’information d’une séquence pour créer une grammaire représentant des éléments

$A_1$	I	II-	$V^7$	I	II-	$V^7$	
	$I^7$	IV <sup>7</sup>	III-	VI <sup>7</sup>	II-	$V^7$	
$A_2$	I	II-	$V^7$	III-	VI <sup>7</sup>	II-	$V^7$
	V-	$I^7$	IV	♯IV <sup>o</sup>	I	I	
$B$	III <sup>7</sup>	III <sup>7</sup>	VI <sup>7</sup>	VI <sup>7</sup>	VI <sup>7</sup>	VI <sup>7</sup>	
	VI-	II <sup>7</sup>	II-	II-	II-	$V^7$	
$A_1$	I	II-	$V^7$	I	VI <sup>7</sup>	II-	$V^7$
	I	$I^7$	IV	♯IV <sup>o</sup>	I	II-	$V^7$

FIGURE 4 – Exemple de progression d’accords générée par la grammaire syntagmatique pour les anatoles.

$B$	$bII^{-7}$	$bV^7$	I <sup>-7</sup>	IV <sup>7</sup>				
	VII <sup>-7</sup>	III <sup>7</sup>	♭VII <sup>-7</sup>	♭III <sup>7</sup>	VI <sup>-7</sup>	II <sup>7</sup>	♭VI <sup>-7</sup>	♭II <sup>7</sup>

FIGURE 5 – Section  $B$  de l’anatole *The Eternal Triangle* de Sonny Stitt.

structurels de la musique. Contrairement aux travaux présentés par Guichaoua [2017] sur l’utilisation de grammaire à dérivation unique [Gallé, 2011] pour la description de structure musicale dans un morceau, nous proposons une méthode probabiliste utilisant un corpus d’apprentissage plus large. Ceci permet de pouvoir déterminer des structures communes dans des scénarios équivalents et permet également d’éviter les problèmes de modélisation lors de l’apparition de variations locales d’un même motif.

La première étape de notre méthode est d’effectuer une segmentation du scénario en unifiant en un symbole les paires de symboles partageant l’information mutuelle la plus forte. Pour deux symboles  $a$  et  $b$  consécutifs, l’information mutuelle  $J(a, b)$  est définie par

$$J(a, b) = \log \frac{\text{count}(a.b)T}{\text{count}(a)\text{count}(b)}, \quad (3)$$

où  $\text{count}$  est la fonction de comptage et  $T$  est la taille du corpus d’apprentissage. Une grande valeur d’information mutuelle signifie que les symboles  $a$  et  $b$  apparaissent consécutivement de manière bien plus fréquente qu’ils ne le seraient par pur hasard. Le regroupement itératif des symboles partageant l’information mutuelle la plus forte permet alors de former une structure hiérarchique du scénario.

Nous avons alors appliqué cette méthode sur le sous-corpus d’anatoles de l’*Omnibook*, constituant 26 variations différentes d’anatoles. Les résultats obtenus sont mitigés et plusieurs limites ont été détectées :

- L’application itérative de cette méthode a tendance à provoquer un phénomène d’agglutination autour des symboles rares, limitant les regroupements des symboles par petits groupes.

- Aucune relation autre que l'identité n'existe entre les différents symboles. Par exemple, toutes les variations de cellules-anatoles sont considérées strictement différentes (I VI- II- V<sup>7</sup> est symboliquement différent de III- VI<sup>7</sup> II- V<sup>7</sup> alors que ces séquences ont la même fonction tonale).
- Bien que l'organisation aux plus bas niveaux semble raisonnable, lorsque l'on considère des niveaux d'organisation plus élevée, les résultats obtenus sont très mauvais. Ceci est lié au problème précédent, car à plus hauts niveaux, nous obtenons un alphabet trop grand pour trop peu de données.

Tout d'abord, pour réduire le premier problème, nous introduisons la notion de durée des symboles, correspondant à la durée des accords qu'ils représentent. Pour un symbole  $a$ , nous notons sa durée  $l(a)$ . Ainsi, pour éviter le problème d'agglutination autour des symboles rares, nous souhaitons privilégier l'unification des paires des symboles de courtes durées. Pour cela, nous proposons une normalisation de l'information mutuelle par la durée des symboles :

$$\tilde{J}(a, b) = \frac{1}{l(a) + l(b)} \log \frac{\text{count}(a.b)T}{\text{count}(a)\text{count}(b)} . \quad (4)$$

Ensuite, pour réduire le second problème, nous proposons lors de la création d'un nouveau symbole de vérifier s'il peut être considéré comme une variation d'un symbole d'une même durée en utilisant la structure séquentielle. Nous souhaitons considérer comme équivalent deux symboles d'une même durée avec des voisinages similaires en terme d'information mutuelle. Deux symboles  $a$  et  $b$  sont considérés équivalents si

$$\Psi(a, b) = \frac{1}{K} \sum_{u,v} (J(u, a) - J(u, b))^2 + (J(a, v) - J(b, v))^2 \leq \xi , \quad (5)$$

avec  $K$  la taille du vocabulaire,  $u$  la liste des symboles à gauche de  $a$  et  $b$ ,  $v$  la liste des symboles à droite de  $a$  et  $b$  et  $\xi$  un seuil à déterminer expérimentalement en fonction du corpus.

Nous proposons alors l'Algorithme 5 pour effectuer l'induction de grammaire à partir d'un corpus de scénario.

---

**Algorithme 5** Induction de grammaire sur un corpus de scenario

---

- Entrée :** Corpus de scénarios.  
**Sortie :** Liste de règles de réécriture.
- 1: **Répéter**
  - 2: Trouver  $a$  et  $b$  tels que  $\tilde{J}(a, b) = \max_{x,y} \tilde{J}(x, y)$ .
  - 3: Créer la règle de réécriture  $X_{ab} \rightarrow a + b$ .
  - 4:  $l(X_{ab}) \leftarrow l(a) + l(b)$ .
  - 5: Remplacer toutes les occurrences de  $a + b$  par  $X_{ab}$  dans le corpus.
  - 6: **Si**  $\exists$  un symbole  $Y$  tel que  $l(Y) = l(X_{ab})$  **et**  $\Psi(Y, X_{ab}) < \xi$  **alors**
  - 7: Créer la règle de réécriture  $Y \rightarrow X_{ab}$ .
  - 8: Remplacer toutes les occurrences de  $X_{ab}$  par  $Y$  dans le corpus.
  - 9: **Fin Si**
- 

Nous avons appliqué cet algorithme sur le sous-corpus d'anatoles de l'*Omni-book*. La Figure 6 montre l'analyse hiérarchique obtenue sur un exemple d'anatole. L'analyse automatique de 10 anatoles a été vérifiée et validée par Pascal Mabit. Premièrement, les différentes fonctions harmoniques (cellules-anatoles, cellules-christophe, etc.) sont retrouvées

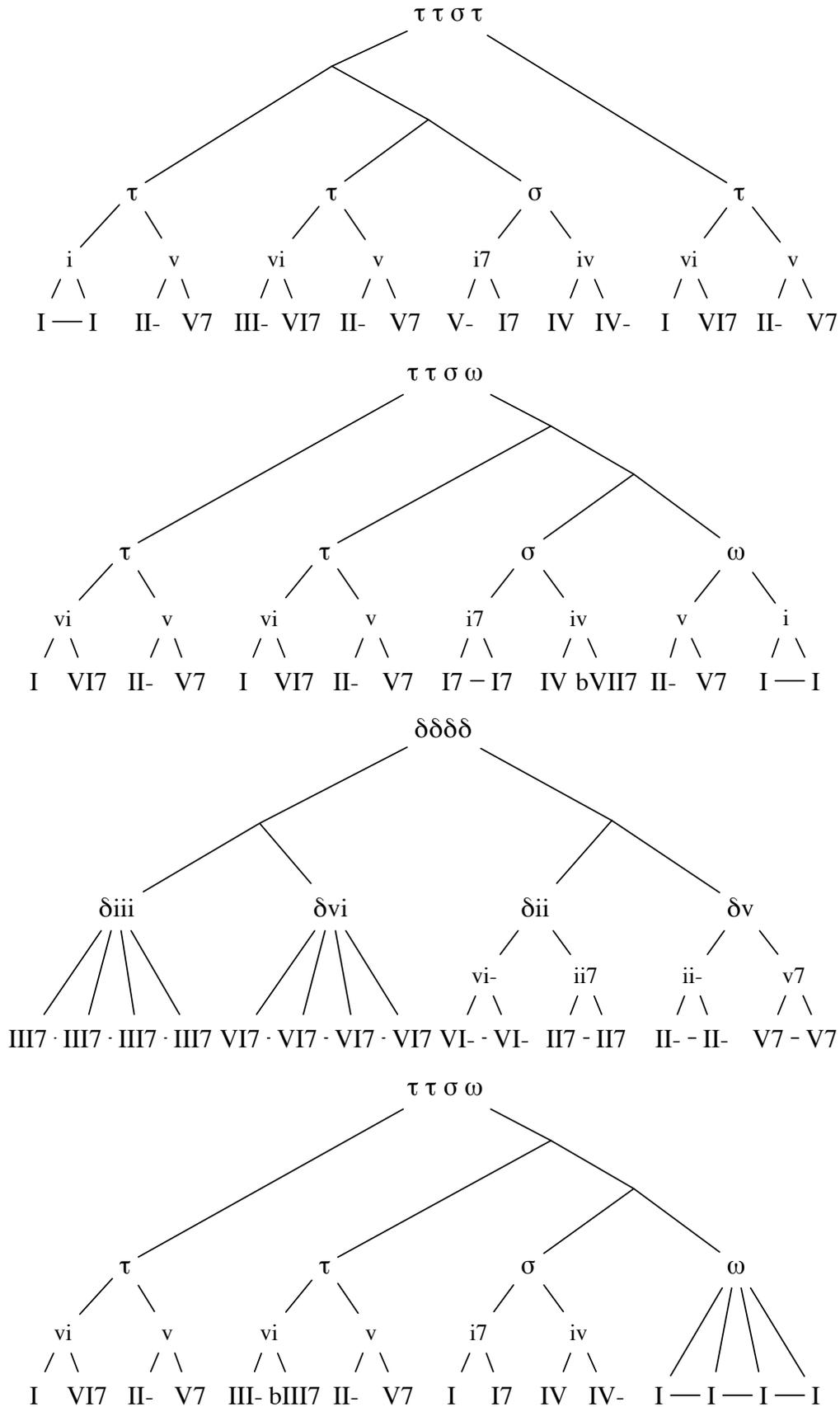


FIGURE 6 – Analyse hiérarchique obtenu à partir de l’Algorithme 4 d’un anatole.  $\mathbb{L}_0$  dernier niveau d’organisation n’est pas représenté ici.

correctement et l'organisation structurelle est correcte aux différentes échelles souhaitées : accords, fonctions tonales et sections. Deuxièmement, les différentes variations d'une même fonction harmonique sont correctement identifiées comme équivalents. Par exemple, I VI<sup>7</sup> II- V<sup>7</sup> et III- bIII<sup>7</sup> II- V<sup>7</sup> sont jugés équivalent, de même pour V- I<sup>7</sup> IV IV- et I<sup>7</sup> I<sup>7</sup> IV bVII<sup>7</sup>.

Cependant, cette analyse est moins précise que celle obtenue par la grammaire créée dans la partie précédente sur certains points :

- Les différents types de sections  $A$  (appelés  $A_1$  et  $A_2$  dans la Grammaire 2) sont considérées comme strictement différentes.
- L'ensemble des variations de cellules-anatoles ont été jugés équivalentes effaçant alors le fait que la première cellule-anatole d'un  $A$  doit commencer par un accord de degré I (ce sous-ensemble de cellule-anatoles est appelé  $\tau_I$  dans la Grammaire 2).

Finalement, les résultats de cette méthode sont encourageants. Les structures multi-niveaux obtenues sont très proches de celles obtenues par la grammaire créée avec un musicien professionnel. Cependant, nous avons utilisé ici un corpus homogène de scénario. Il serait intéressant de voir si des résultats satisfaisants pourraient être obtenus sur des corpus plus larges, mais aussi plus variés, comme par exemple celui du *Realbook*.

### 3 Exploiter une structure multi-niveaux dans l'improvisation guidée

Dans cette partie, nous présentons comment utiliser l'ensemble de l'information d'un scénario multi-niveaux pour enrichir la méthode d'improvisation de scénario temporel d'*ImproteK* présentée par Nika et al. [2017] en prenant en compte la structure hiérarchique du scénario fournie par la grammaire syntagmatique. Nous rappelons d'abord les principes utilisés par Nika et al. pour générer des improvisations sur un scénario, puis nous proposons une méthode pour utiliser l'information multi-niveaux pour étendre les possibilités de génération.

#### 3.1 Improvisation sur un scénario temporel

L'idée principale d'*ImproteK* est d'utiliser des connaissances *a priori* d'un scénario temporel afin d'introduire des mouvements d'anticipation dans le guidage de la génération d'improvisations. Dans le cas d'*ImproteK*, les scénarios sont représentés par une séquence de symboles appelés *étiquettes*. Contrairement à *OMax* et aux modèles que nous avons présentés dans Déguernel et al. [2016, 2018], au cours de l'apprentissage, la mémoire du système n'est pas organisée selon la dimension qui est improvisée, mais est une séquence de contenus musicaux de la dimension improvisée organisée à partir des étiquettes de scénario. Pour reprendre l'exemple du jazz, les étiquettes peuvent être les accords d'une progression harmonique et les contenus les notes de la mélodie jouée par un improvisateur. Lors du processus de génération, pour un scénario donné, on cherche alors à combiner par rapport aux étiquettes une anticipation des éléments du scénario futur avec une navigation de la mémoire de manière similaire à *OMax* pour conserver une cohérence stylistique. Les contenus musicaux des états choisis sont alors joués pour générer l'improvisation.

Le processus de génération d'une improvisation se déroule alors en deux étapes successives d'anticipation puis de navigation (voir [Nika et al., 2017] pour une description

plus détaillée des algorithmes). Notons  $S = S_1 \dots S_s$  le scénario,  $T$  la position courante du scénario et considérons une mémoire construite avec un oracle des facteurs constitué des états  $0 \dots m$  avec les étiquettes  $\Lambda_0 \dots \Lambda_m$ .

1. L'étape d'anticipation consiste à chercher des événements dans la mémoire partageant un futur commun avec la position courante du scénario tout en garantissant une continuité avec le passé de la mémoire. Cette étape est effectuée en indexant dans la mémoire les préfixes des suffixes de la position courante du scénario. On construit d'abord l'ensemble des états de la mémoire partageant un futur commun avec la position courante du scénario  $S_T \dots S_s$  :

$$\text{Futur}(T) = \{j \in [0 \dots m] \mid \exists c_{\text{futur}} \in \mathbb{N}, \Lambda_j \dots \Lambda_{j+c_{\text{futur}}-1} \in \text{Pref}(S_T \dots S_s)\} . \quad (6)$$

Les  $j \in \text{Futur}(T)$  sont alors les positions de début des facteurs de la mémoire correspondant à un préfixe du scénario à partir de la position courante. La taille de ces facteurs est mesurée par  $c_{\text{futur}}$ .

Puis, on construit l'ensemble des états de la mémoire partageant un passé commun avec l'état courant  $i$  de la mémoire :

$$\text{Passé}(i) = \{j \in [0 \dots m] \mid \exists c_{\text{passé}} \in [1, j], \Lambda_{j-c_{\text{passé}}+1} \dots \Lambda_j \in \text{Suff}(0 \dots i)\} . \quad (7)$$

Les  $j \in \text{Passé}(i)$  sont les positions de fin des facteurs de la mémoire correspondant à un suffixe de l'évènement courant dans la mémoire. La taille de ces facteurs est mesurée par  $c_{\text{passé}}$ . La construction de cet ensemble est en pratique effectuée en utilisant les propriétés sur les liens suffixiels de l'oracle des facteurs.

Pour l'étape d'anticipation, on cherche alors les positions  $j$  dans la mémoire telles que :

$$\text{Ant}(T, i) = \{j \in [0 \dots m] \mid j \in \text{Futur}(T) \wedge j - 1 \in \text{Passé}(i)\} . \quad (8)$$

La Figure 7 illustre cette étape d'anticipation.

2. L'étape de navigation consiste à chercher des événements de la mémoire partageant un contexte commun avec la position courante du scénario tout en respectant le scénario. On cherche les positions  $j$  dans la mémoire telles que :

$$\text{Nav}(T, i) = \{j \in [0 \dots m] \mid \Lambda_j = S_T \wedge j - 1 \in \text{Passé}(i)\} . \quad (9)$$

Cette étape permet alors d'effectuer des chemins non-linéaires dans la mémoire permettant de créer des nouvelles phrases musicales, de manière similaire à *OMax* [Assayag & Bloch, 2007] ou des modèles présentés par Déguernel et al. [2016]. Cela permet au systèmes des variations plus locales que si l'on utilisait l'étape d'anticipation seule.

En pratique des valeurs minimale et maximale pour  $c_{\text{futur}}$  et  $c_{\text{passé}}$  sont fixées pour éviter l'utilisation de fragments de la mémoire trop courts ou trop longs lors de la génération de l'improvisation.

### 3.2 Utiliser l'information multi-niveaux

Nous voulons à présent prendre en considération un scénario multi-niveaux pour la génération d'improvisations. Nous nous mettons alors dans le cas où le scénario n'est plus une simple séquence de symboles, mais une séquence de listes de symboles correspondant à

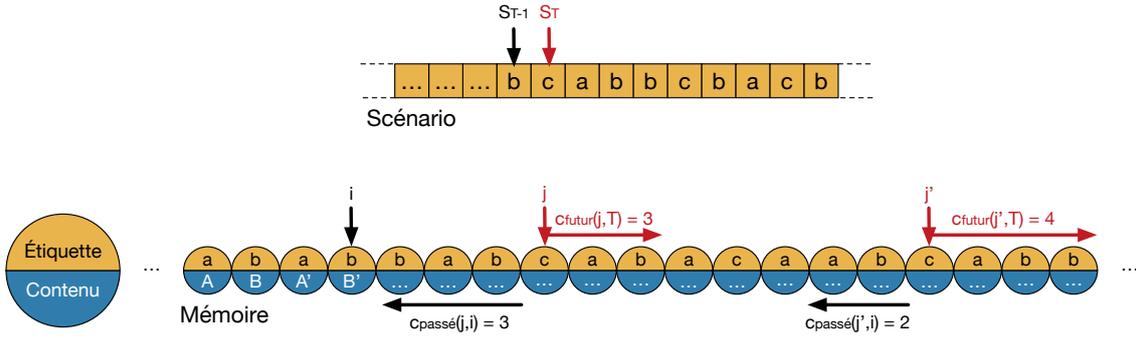


FIGURE 7 – Construction de l'ensemble des éléments de la mémoire partageant un futur commun avec la position courante du scénario  $T$  et un passé commun avec l'évènement courant de la mémoire  $i_{T-1}$  pour l'étape d'anticipation du processus de génération d'improvisation sur un scénario temporel [Nika et al., 2017].

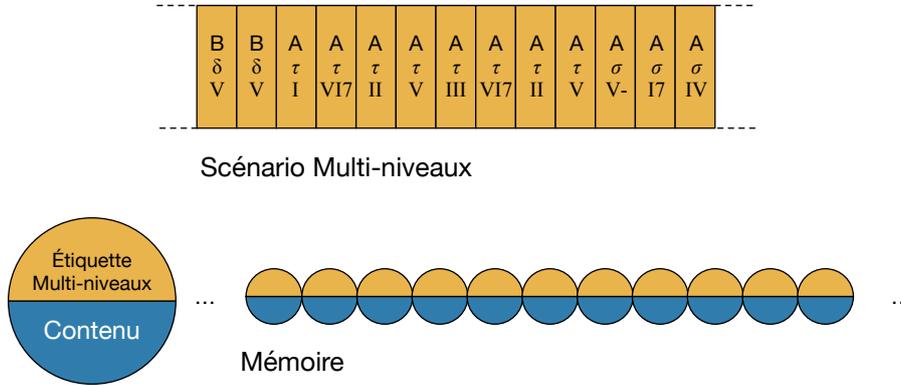


FIGURE 8 – Représentation d'un scénario et d'une mémoire multi-niveaux.

chaque niveau. Pendant la construction de la mémoire, celle-ci est alors organisée à partir d'étiquettes multi-niveaux correspondant aux scénarios multi-niveaux. La Figure 8 montre un exemple d'un scénario multi-niveaux avec trois niveaux d'organisation temporelle.

Nous adaptons alors la méthode de génération présentée dans la partie précédente pour prendre en compte l'information ajoutée par l'aspect multi-niveaux à la fois pour l'étape d'anticipation et pour l'étape de navigation. Pour chaque étape nous souhaitons étendre les possibilités de positions dans la mémoire à des états partageant des étiquettes multi-niveaux équivalentes à celles du scénario, c'est-à-dire des étiquettes partageant des informations communes à certains niveaux mais différents à d'autres niveaux. Par exemple, dans le cas du jazz, nous pouvons vouloir accepter des positions dans la mémoire qui n'ont pas le même accord que celui dans le scénario, mais qui possèdent la même fonction tonale et la même section. De cette manière, il est possible de générer des improvisations sur un scénario avec un accord n'existant pas dans la mémoire, mais partageant la même fonction que des accords connus. Nous voulons également être capables de favoriser les positions de la mémoire partageant une information commune forte avec le futur du scénario et le passé de la mémoire. Ainsi la génération prend mieux compte de la structure hiérarchique du scénario pour guider l'improvisation.

Ce processus peut se formaliser comme suit. Notons  $S = S_1 \dots S_s$  le scénario multi-niveaux de longueur  $s$  avec  $Q$  niveaux, avec  $\forall p \in [1 \dots S], S_p = \{S_p^1 \dots S_p^Q\}$ . Soit  $T$  la position courante du scénario et considérons une mémoire construite avec un oracle des facteurs constitué des états  $0 \dots m$  avec les étiquettes multi-niveaux  $\Lambda_0 \dots \Lambda_m$ , avec  $\forall p \in [0 \dots m], \Lambda_p = \{\Lambda_p^1 \dots \Lambda_p^Q\}$ .

On dit qu'une séquence de symboles multi-niveaux  $u = u_1 \dots u_n$  avec  $\forall p \in [1 \dots n], u_p = \{u_p^1 \dots u_p^Q\}$  est un préfixe équivalent à une séquence de symboles multi-niveaux  $v = v_1 \dots v_m$  avec  $\forall p \in [1 \dots m], v_p = \{v_p^1 \dots v_p^Q\}$  par :

$$u = u_1 \dots u_n \in \text{Pref\_eq}(v = v_1 \dots v_m) \text{ ssi } \forall p \in [1 \dots n], \exists q; u_p^q = v_p^q . \quad (10)$$

De même, on dit qu'une séquence de symboles multi-niveaux  $u = u_1 \dots u_n$  avec  $\forall p \in [1 \dots n], u_p = \{u_p^1 \dots u_p^Q\}$  est un suffixe équivalent à une séquence de symboles multi-niveaux  $v = v_1 \dots v_m$  avec  $\forall p \in [1 \dots m], v_p = \{v_p^1 \dots v_p^Q\}$  par :

$$u = u_1 \dots u_n \in \text{Suff\_eq}(v = v_1 \dots v_m) \text{ ssi } \forall p \in [1 \dots n], \exists q; u_p^q = v_{p+m-n}^q . \quad (11)$$

Les deux étapes de génération sont alors modifiées pour prendre en compte l'information multi-niveaux et les étiquettes équivalentes.

1. Pour l'étape d'anticipation, on remplace l'ensemble  $\text{Futur}(T)$  par l'ensemble  $\text{Futur\_eq}(T)$  défini par :

$$\text{Futur\_eq}(T) = \{j \in [0 \dots m] \mid \exists c_{\text{futur}} \in \mathbb{N}, \Lambda_j \dots \Lambda_{j+c_{\text{futur}}-1} \in \text{Pref\_eq}(S_T \dots S_s)\} . \quad (12)$$

Les  $j \in \text{Futur\_eq}(T)$  sont les positions de début des facteurs de la mémoire correspondant à un préfixe équivalent du scénario à partir de la position courante.

On remplace également l'ensemble  $\text{Passé}(i)$  par l'ensemble  $\text{Passé\_eq}(i)$  défini par :

$$\text{Passé\_eq}(i) = \{j \in [0 \dots m] \mid \exists c_{\text{passé}} \in [1, j], \Lambda_{j-c_{\text{passé}}+1} \dots \Lambda_j \in \text{Suff\_eq}(0 \dots i)\} . \quad (13)$$

Les  $j \in \text{Passé\_eq}(i)$  sont les positions de fin des facteurs de la mémoire correspondant à un suffixe équivalent de l'évènement courant dans la mémoire.

Finalement, pour l'étape d'anticipation, on considère toutes les positions  $j$  dans la mémoire telles que :

$$\text{Ant\_eq}(T, i) = \{j \in [0 \dots m] \mid j \in \text{Futur\_eq}(T) \wedge j - 1 \in \text{Passé\_eq}(i)\} . \quad (14)$$

2. Les modifications pour l'étape de navigation sont similaires ; on étend les possibilités grâce aux étiquettes équivalentes. Ainsi, on cherche les positions  $j$  dans la mémoire telles que :

$$\text{Nav\_eq}(T, i) = \{j \in [0 \dots m] \mid \Lambda_j \sim S_T \wedge j - 1 \in \text{Passé\_eq}(i)\} \quad (15)$$

où  $\Lambda_j \sim S_T$  si et seulement si  $\exists q; \Lambda_j^q = S_T^q$ .

Comme dans la partie précédente, en pratique des valeurs minimale et maximale pour  $c_{\text{futur}}$  et  $c_{\text{passé}}$  sont fixées pour éviter l'utilisation de fragments de la mémoire trop courts ou trop longs lors de la génération de l'improvisation.

Afin de privilégier à chaque étape les positions  $j$  partageant le plus d'information commune avec le futur du scénario et le passé de la mémoire, un score de similarité entre étiquettes multi-niveaux est défini. L'utilisateur peut attribuer *a priori* à chaque niveau temporel  $q$  considéré un poids  $W_q$  tel que

$$\sum_q W_q = 1 . \quad (16)$$

Par exemple, dans le cas du jazz, on peut considérer que le niveau de la fonction tonale est plus important que celui des accords ou que celui des sections.

La similarité entre deux étiquettes multi-niveaux  $\Lambda_i$  et  $\Lambda_j$  est alors égale à

$$\varsigma(\Lambda_i, \Lambda_j) = \sum_q W_q \delta_{\Lambda_i^q, \Lambda_j^q} , \quad (17)$$

où  $\delta$  est le symbole de Kronecker.

Chaque élément  $j$  dans  $\text{Ant\_eq}(T, i)$  peut alors obtenir un score défini par :

$$\sum_{k=0}^{c_{\text{futur}}(j, T)} \varsigma(\Lambda_{j+k}, S_{T+k}) + \sum_{k=0}^{c_{\text{passé}}(j, i)} \varsigma(\Lambda_{j-k}, \Lambda_{i-k}) \quad (18)$$

et chaque élément  $j$  dans  $\text{Nav\_eq}(T, i)$  peut obtenir un score défini par :

$$\varsigma(\Lambda_j, S_T) + \sum_{k=0}^{c_{\text{passé}}(j, i)} \varsigma(\Lambda_{j-k}, \Lambda_{i-k}) . \quad (19)$$

À partir de ces scores, différentes stratégies peuvent être appliquées pour privilégier les éléments de  $\text{Ant\_eq}(T, i)$  et  $\text{Nav\_eq}(T, i)$  partageant une information commune forte avec le futur du scénario et le passé de la mémoire. L'élément avec le plus haut score peut être choisi, ou on peut limiter le choix aux éléments avec un score supérieur à un certain seuil. Dans nos expériences, nous normaliserons ces scores par la somme des scores de tous les évènements possibles et nous choisirons un élément de manière aléatoire en suivant les probabilités obtenues.

## 4 Évaluation et retours des musiciens

### 4.1 Choix des paramètres

Pour évaluer l'utilisation de scénarios multi-niveaux pour la génération d'improvisation, nous avons réalisé des tests d'écoute avec des musiciens professionnels afin d'obtenir une évaluation qualitative des générations. Nous avons fait appel à deux musiciens et improvisateurs professionnels :

- Louis Bourhis, contrebassiste de jazz, diplômé de la Haute École de Musique de Lausanne,

— Joël Gauvrit, pianiste et professeur de jazz et de musique classique, diplômé du Conservatoire National Supérieur de Musique et de Danse de Lyon.<sup>2</sup>

Les tests d'écoute ont été faits séparément.

Nous nous sommes intéressés au cas des anatoles issus du corpus de l'*Omnibook* de Charlie Parker. Nous avons tout d'abord utilisé ce sous-corpus pour apprendre les symboles terminaux de la grammaire syntagmatique présentée dans la partie 1.2. Cette grammaire a ensuite été utilisée pour générer des scénarios multi-niveaux d'anatoles avec trois niveaux d'organisation : les accords, les fonctions tonales et les sections. Sur ces scénarios, nous avons généré des improvisations en suivant deux méthodes :

1. La méthode de génération basée sur *ImproteK* présentée dans la partie 3.1. L'anticipation du scénario et la navigation dans la mémoire sont alors effectuées ici seulement sur le niveau des accords du scénario (les autres niveaux ne sont pas pris en compte). Lorsqu'un accord de scénario n'apparaît pas dans la mémoire, des principes de transposition sont appliqués.
2. La méthode de génération basée sur un scénario multi-niveaux présentée dans la partie 3.2. L'anticipation du scénario et la navigation dans la mémoire sont effectuées avec l'ensemble de l'information multi-niveaux. Nous avons considéré ici, avec l'appui d'un musicien, que pour les anatoles, le niveau le plus important est le niveau des fonctions tonales et nous avons attribué un poids de 0,5 à ce niveau. Puis nous avons attribué un poids de 0,3 au niveau des accords et un poids de 0,2 au niveau des sections.

Dans les deux cas, les contenus de la mémoire sont basés sur un thème et improvisation issu d'un anatole de l'*Omnibook*. Contrairement aux expériences précédentes, le découpage dans la mémoire n'est pas effectué note-à-note mais temps-à-temps pour pouvoir faire correspondre plus facilement les étiquettes de la mémoire avec les étiquettes du scénario. La Figure 9 résume le processus complet de génération utilisé dans le cas multi-niveau. Des exemples d'improvisations générées par les deux méthodes sont disponibles sur : [repmus.ircam.fr/dyci2/ressources](http://repmus.ircam.fr/dyci2/ressources). Les Figures 11 et 12 montrent également des relevés d'improvisations générées dans le cas multi-niveaux respectivement avec la mémoire apprise sur *Thriving from a Riff* et sur *An Oscar for Treadwell*.

## 4.2 Résultats et analyses

Chaque musicien a écouté au total une douzaine d'improvisations, plus ou moins, selon son vouloir. À chaque fois, pour une même progression d'accords générée par la grammaire syntagmatique, une improvisation était générée par chacune des deux méthodes. La mémoire de l'oracle des facteurs a été construite à partir de trois morceaux issus du sous-corpus d'anatoles de l'*Omnibook* : *Thriving from a Riff*, *An Oscar for Treadwell*, et *Anthropology*.

Dès les premiers exemples, les deux musiciens ont remarqué une différence très nette dans la génération des deux méthodes. L'amélioration majeure constatée par les musiciens est la construction globale de l'improvisation sur l'ensemble de la grille. Les improvisations générées par la deuxième méthode fournissent des résultats jugés plus réalistes grâce à une meilleure prise en compte des structures harmoniques. Gauvrit dit :

« C'est incroyable la différence. Ah ouais ! Bah oui là on a l'impression qu'il construit. Ah bah là c'est super. C'est un super chorus. J'ai un élève qui fait ça je

2. Des biographies plus détaillées des musiciens sont fournies en Annexe B.

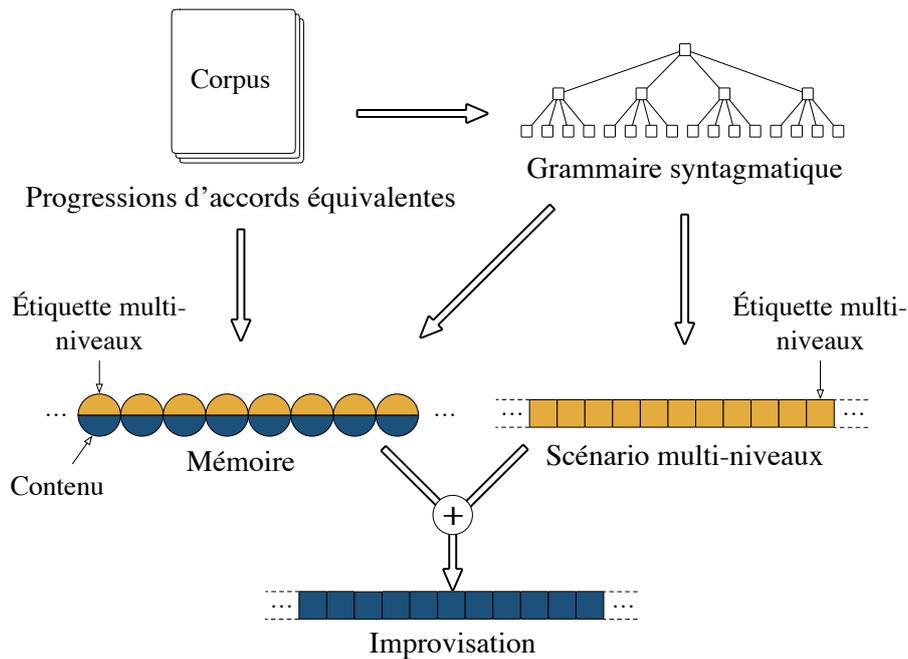


FIGURE 9 – Processus de génération d'improvisation sur un scénario multi-niveaux. Tout d'abord, une grammaire syntagmatique est construite avec un musicien à partir d'un corpus, puis les symboles terminaux de cette grammaire sont appris sur ce corpus. Cette grammaire est utilisée pour générer des scénarios multi-niveaux et pour fournir les informations multi-niveaux lors de l'apprentissage de la mémoire avec des étiquettes multi-niveaux. Finalement, les improvisations sont générées en anticipant le scénario multi-niveaux et en navigant la mémoire avec les étiquettes multi-niveaux.

suis content. C'est-à-dire que même rythmiquement, la façon dont il fonctionne à la structure harmonique. J'entends ça moi. Que l'autre non, du tout. »

Bourhis précise cette idée en faisant remarquer qu'avec la deuxième méthode, l'improvisation s'axe autour des fonctions et des sections, formant des phrases (ou ensembles de phrases) avec une logique à plus long terme. Cela peut être vu en particulier sur la Figure 11 :

« L'architecture globale est beaucoup plus réfléchie. On sent qu'il est beaucoup plus à l'aise. Enfin l'ordinateur n'est pas à l'aise, mais nous on a l'impression d'écouter quelqu'un qui est beaucoup plus avancé. [...] C'est-à-dire, quand il commence une partie, il commence avec moins de notes et il va fournir dessus. Je trouve que c'est plus humain dans ce sens-là. »

Les mélodies des improvisations générées par la deuxième méthode ont également été jugées généralement moins décousues. Cela peut s'expliquer par le fait que la deuxième méthode ouvre des possibilités dans la navigation permettant de faire moins de sauts dans la mémoire lors de progressions harmoniques locales difficiles. Bourhis dit :

« C'est vrai qu'il y a les trucs, tu sais, du tout début où on a un peu peur, que l'ordi il prenne ce fragment, ce fragment et ce fragment et qu'il les pose les uns à la suite des autres sans trop les relier. Et du coup, ce que je t'ai expliqué tout



quement. J'ai l'impression de voir des stades différents chez mes élèves de compréhension tu vois, de ce qui est écrit, analytique et tout... Et c'est marrant la perception spontanément, parce qu'après quand t'y penses, c'est naturel, ce que ça engendre rythmiquement, le fait d'avoir conscience des structures harmoniques, ça engendre rythmiquement des choses qui sont beaucoup plus naturelles. »

Généralement, la deuxième méthode a été très peu critiquée par les musiciens. Bourhis a noté encore l'existence de sauts dans la mémoire un peu perturbants (avec par exemple des sauts d'octave peu réalistes), mais leur fréquence est bien moindre avec la deuxième méthode qu'avec la première :

« Après, il y a toujours tous ces trucs tordus, mais tout est renforcé avec la première méthode quand c'est décousu, les sauts d'octaves principalement. »

Finalement, Gauvrit étend par la suite l'intérêt pédagogique que ces différentes méthodes peuvent illustrer, en ré-explicant que l'intégration des structures harmoniques dans le processus de génération apporte un réalisme notable :

« Et moi j'y vois une application pédagogique géniale, parce que tu montres ça à des élèves. Je pense que c'est absolument frappant. Puisque finalement, ça correspond à la réalité... Et puis c'est marrant de voir que le stade... Enfin moi je trouve que le déclic du moment où vraiment on y croit, c'est au moment où il y a une compréhension analytique. C'est là où on s'en rend compte... Ça sert ma crèmerie, c'est pour ça que ça m'intéresse. »

### 4.3 Résumé des résultats

Les retours des musiciens sur cette expérience sont très satisfaisants. La prise en considération de la structure multi-niveaux d'une progression harmonique permet la génération d'improvisations beaucoup plus réalistes. En plus de pouvoir improviser de manière plus libre et varier (même sur des accords non vus auparavant), celle-ci permet une amélioration de la cohérence mélodique et rythmique. Généralement, la génération « ne se préoccupe pas trop du détail au service de ce qui peut se passer à plus grande échelle », se rapprochant alors des processus d'improvisation employés par des musiciens humains, permettant une meilleure considération des espaces harmoniques et un meilleur phrasé.

The musical score is written in 4/4 time and consists of eight staves. The key signature has two flats (Bb and Eb). The chords used are: Bb, C-, F7, Bb, C-, F7 (Staff 1); Bb7, Eb7, D-, G7, C-, F7 (Staff 2); Bb, C-, F7, D-, G7, C-, F7 (Staff 3); F-, Bb7, Eb, E<sup>o</sup>, Bb, Bb (Staff 4); D7, D7, G7, G7 (Staff 5); G-, C7, C-, F7 (Staff 6); Bb, C-, F7, Bb, G7, C-, F7 (Staff 7); Bb, Bb7, Eb, E<sup>o</sup>, Bb, C-, F7 (Staff 8). The melody includes several triplet markings (3).

FIGURE 11 – Improvisation avec la mémoire de *Thriving from a Riff*. Progression d'accords générée avec la grammaire syntagmatique pour anatoles.

The musical score is written in 4/4 time and consists of eight staves of music. The chords and melodic lines are as follows:

- Staff 1: Chords C, D-, G7, E-, A7, D-, G7. Melody includes a triplet of eighth notes.
- Staff 2 (labeled 6): Chords G-, C7, F7, F-, C, A7, D-, G7. Melody includes a triplet of eighth notes.
- Staff 3 (labeled 10): Chords C, A7, D-, G7, C, A7, D-, G7. Melody includes a triplet of eighth notes.
- Staff 4 (labeled 14): Chords C7, F7, Bb7, D-, G7, C. Melody includes two triplet markings.
- Staff 5 (labeled 18): Chords B-, E7, E-, A7. Melody includes a triplet of eighth notes.
- Staff 6 (labeled 22): Chords A-, D7, D-, G7. Melody includes a triplet of eighth notes.
- Staff 7 (labeled 26): Chords C, A7, D-, G7, E-, A7, D-, G7. Melody includes a triplet of eighth notes.
- Staff 8 (labeled 30): Chords C, C7, F7, F-, C, C. Melody includes a triplet of eighth notes.

FIGURE 12 – Improvisation avec la mémoire de *An Oscar For Treadwell*. Progression d'accords générée avec la grammaire syntagmatique pour anatoles.

# Bibliographie

- G erard Assayag & Georges Bloch (2007). Navigating the oracle : a heuristic approach. In *Proceedings of the 33rd International Computer Music Conference* (pp. 405–412).
- Noam Chomsky (1956). Three models for the description of language. *IEEE Transactions on Information Theory*, 3(2), 113–124.
- Noam Chomsky (1972). *Studies on semantics in generative grammar*. Walter de Gruyter.
- Carl de Marcken (2015). Lexical heads, phrase structure and the induction of grammar. In *Proceedings of the 3rd Workshop on Very Large Corpora* (pp. 14–26).
- Ken D eguernel (2018). *Apprentissage de structures musicales en contexte d'improvisation*. PhD thesis, Universit e de Lorraine.
- Ken D eguernel, J er ome Nika, Emmanuel Vincent, & G erard Assayag (2017). Generating equivalent chord progressions to enrich guided improvisation : application to rhythm changes. In *Proceedings of the the 14th Sound and Music Computing Conference* (pp. 399–406).
- Ken D eguernel, Emmanuel Vincent, & G erard Assayag (2016). Using multidimensional sequences for improvisation in the OMax paradigm. In *Proceedings of the 13th Sound and Music Computing Conference* (pp. 117–122).
- Ken D eguernel, Emmanuel Vincent, & G erard Assayag (accept e 2018). Probabilistic factor oracles for multidimensional machine improvisation. *Computer Music Journal*.
- Matthias Gall e (2011). *Searching for compact hierarchical structures in DNA by means of the smallest grammar problem*. PhD thesis, Universit e Rennes 1.
- Mathieu Giraud, Ken D eguernel, & Emiliios Cambouroupoulos (2014). Fragmentations with pitch, rhythm and parallelism constraints for variation matching. In *Proceedings of the 10th International Symposium of Computer Music Multidisciplinary Research* (pp. 298–312).
- Corentin Guichaoua (2017). *Mod eles de compression et crit eres de complexit e pour la description et l'inf erence de structure musicale*. PhD thesis, Universit e Rennes 1.
- John E. Hopcroft & Jeffrey D. Ullman (1979). *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley.
- J er ome Nika (2016). *Guiding human-computer music improvisation : introducing authoring and control with temporal scenarios*. PhD thesis, UPMC - Universit e Paris 6 Pierre et Marie Curie.
- J er ome Nika, Marc Chemillier, & G erard Assayag (2017). ImproteK : introducing scenarios into human-computer music improvisation. *ACM Computers in Entertainment*, 4(2), 4 :1–27.

---

John K. Pate & Mark Johnson (2016). Grammar induction from (lots of) words alone. In *Proceedings of the 26th International Conference on Computational Linguistics* (pp. 23–32).

Jacques Siron (2015). *La partition intérieure – jazz, musiques improvisées. 9ème édition revue et corrigée*. Outre mesure.

Imed Zitouni, Kamel Smaïli, & Jean-Paul Haton (2000). Beyond the conventional statistical language models : the variable-length sequences approach. In *Interspeech* (pp. 562–565).