

## 研究報告

## ジャズ単音アドリブソロにおける音高認知の調査 Case-Studies about Pitch Recognition of Jazz Ad-lib Solo Phrases

安藤 大地

Daichi ANDO

首都大学東京システムデザイン学部インダストリアルアート学科

Division of Industrial-Art, Faculty of System Design

Tokyo Metropolitan University

## 概要

本稿では、何人かのジャズプロ演奏家と学習者のジャズアドリブソロにおける音高認知についての事例調査の報告を行う。D.Sudnow のジャズアドリブ習得過程の一人称的研究や、著者が過去に行ったジャズベーシストへの調査から、フレーズの音高認知はその楽器の物理的制約に依存する傾向が強いと著者は考えていたが、半年ほどジャズサクソ奏者（プロ演奏家）、ジャズギター奏者（プロ演奏家とその演奏家に師事している学習者）にそれぞれ聞き取り調査を行い、さらに著者自身（ジャズサクソ学習者）の一人称的記録をまとめた結果、同じ楽器でも個人によってさらに音高把握の感覚が大きく異なっていることがわかった。今回調査したジャズサクソプロ演奏家は音高把握を基本的に移動ド・音程で行っており、ジャズサクソ学習者である著者自身は固定ド・「音名を歌う」ことで行っていた。同様に今回調査対象となったジャズギタープロ演奏家は単音フレーズでも指板を抑えるコードフォームで音高を把握しているらしいことが聞き取り調査の結果わかった。

In this paper, the author reports that the jazz performer's perception of ad-lib solo generation. The survey is conducted in three ways: first-person description of the learner, second-person description from the person receiving the lesson, and interview as a third-person of view. The survey target is both the jazz saxophone and the jazz guitar respectively. Based on the results, the author considered that not only the physical characteristics of the instruments but also the embodied knowledge of the individual performers influence the recognition of phrases and pitches.

## 1. はじめに

本研究の最終目的は、人間と音楽知識・音楽意図を共有し協調演奏するジャズ即興演奏 AI の開発である。近年音響ベースの深層学習を用い音響で表現することで少しずつ実現されているが (Esling, Chemla-Romeu-Santos, and Bitton 2018), 音響の生成であり、かつ深層学習の問題であるブラックボックス化が行われていることから、学習した人工知能がどのように演奏を生成しているのかを人間がうかがい知ることができない。

そこで本研究では、人間と人工知能が「音楽的な知識を共有すること」を重要視している。そのために音楽的ロジックとして記述することができる即興演奏の知識獲得を目的とし、人間のジャズ即興演奏のモデル化を、実際のジャズ演奏家の認知を元に進めている。

Sudnow は自身のジャズピアノ即興演奏習得過程を自身の感覚に基づいて記録・記述する研究を行なった (Sudnow 1978)。Sudnow の記述では五線譜は登場せず、鍵盤とそれに対応する手の動きという2つ要素の関係性のみで自らの即興演奏認知が構成されていることを示している。特に和音に適したスケールに関しては白鍵と黒鍵という上下に配置された鍵盤に対する手の動きが具体的に記述されており、少なくとも Sudnow 自身の学習過程においては五線譜もしくは半音刻みのノートナンバー等の音楽情報的なモデルの意識がなかったと推測することができる。

著者はジャズの4ビートベースアドリブバックキングのルールベース生成システムの構築のために、ジャムセッション会場などでアマチュアベーシストに聞き取り調査を行なった (安藤大地 2015)。聞き取りを行なったアマチュアベーシストの多くは、曲のリードシートから読み取っているのは「コードの基音の音名」と「コードの種類」のみであり、それを頭の中で五線譜などの音楽情報媒体にマッピングすることなく、コードの基音とコードの種類を左手の動かし方の感覚で把握していることがわかった。さらに4ビートランニングベー

スラインを構築するための7度やそれを超える音程の跳躍においても、ほぼ手の感覚のみ、すなわち身体知として4ビートバックフレーズを生成していることがわかった。

これら2つの事例から、ジャズ演奏家の即興演奏においては、楽器の物理的な楽器と身体的な特性に基づき、フレーズひいてはフレーズを構成している音高を認知している傾向が強いではないか、つまり同じ楽器の演奏家ならば同じような音高の認知をしているのではないか、という仮説を立てた。

## 2. これまでの調査

### 2.1. 一人称視点での練習と認知の記録と認知

著者は10年以上クラシック音楽でサックス学習と演奏を行っており、2016年よりジャズサックスのプロ演奏家に師事してジャズ即興演奏を学習を始めた。レッスンは月1回もしくは2回だが、毎日もしくは1日おきに、1日30分～1時間程度の練習時間を確保し「どのような練習をした時に、どのような技能を獲得したか、獲得感を得られたか」などを中心に、ジャズ即興演奏習得過程の一人称的な記録を行ってきた。この記録は著者が作曲や楽理などの専門家であり、音楽大学の演奏専攻に在籍経験もあることから、音楽的経験・知識に基づいた記述になっており、Sundnowの練習記録に比べて、

1. 五線譜を多用した上で理解している
2. 体の動かし方を音楽的に理解している

などの特徴がある。

### 2.2. 二人称視点での聞き取り調査

#### 2.2.1. 自身の教授者への聞き取り調査

著者はレッスンを受けると同時に、著者にジャズサックス即興演奏を指導しているジャズサックス演奏家が、著者に師事した練習や手法に対して演奏家自身がどのように認識しているかを聞き取り調査している。

#### 2.2.2. 第三者の学習者と教授者への聞き取り調査

また2019年より、これまでロックギターを演奏して25歳を被験者とし、ジャズギターのプロ演奏家にレッスンを受けてもらい、著者の場合と同じように一人称的に記録されたレッスンや練習の記録を収集するとともに、教授しているプロ演奏家に対する聞き取り調査を行なっている。レッスンは月2回、聞き取り調査は月1回、継続して行なっている。レッスンを受けてもらう被験者には音楽的知識はあまりないため、記

述はそれほど詳しくはなく、著者から質問をするとともに、レッスンを録音したのも合わせて分析している。対してプロ演奏家に対する聞き取り調査では、音楽的経験・知識に基づいた会話で行なっている。

先に立てた仮説を検証するため、著者自身、著者を指導しているジャズサックス演奏家、ジャズギター学習者、指導しているジャズギター演奏家に対して、それぞれジャズ即興を生成するときの音高を認知に関して、質問や記録を元に考察を行なった。

## 3. 現段階でのまとめ

現在までの記録や聞き取り調査から推測されることを以下にまとめる。

### 3.1. ジャズサックス学習者（一人称的記述）

サックスは初めて12年程度。ただし音大の音楽専攻に在学経験あり。

- 固定ドで、階名を頭に浮かべる（頭の中で階名で歌う）ことで指が自動的に動く。またフレーズを聴いた時にそれを再現するためには、聴くフレーズが音高=階名で聴こえている（聴いた時に階名が頭に浮かび、それを覚える）必要がある。
- 固定ドではあるが、音楽専攻であったことから移動ド感覚もある。特に和声短音階、旋律短音階については非常に自然に使用することができる。固定ドの階名で全ての長調・短調を歌うことができる。
- コード構成音についてはピアノ鍵盤上もしくは五線譜情報として展開してからサックスの指の動きに再マッピングすることで習得した。現在はコード構成音はサックスの指の動きとして獲得している。（単音コードアルペジオならば自動的に展開できる）
- 長調・短調のスケール感覚は身につけており、特に旋律的短音階、和声短音階については自然に指が動くなど獲得できている。一方それ以外のスケール感覚、特に人工スケールに関しては把握が難しい（音名が頭に浮かばない）。
- 人工スケールなどのアウト感などについては、クラシック音楽的に五線譜情報で把握している。
- 和音を単音アルペジオとしてしか認識できない。

### 3.2. ジャズサックスプロ演奏家（教授を通じた記述）

（現段階では、本人の認知が明らかになっていない、明確に認知していない部分も多いため、今後さらなる聞き取り調査が必要である。）

テナーサクスがメインでありアルト持ち替えも容易。ピアノ経験はそれほどなし。

- 移動ドの傾向が強い。音程 (Pitch-interval, インターバル) を感じることで指が動く。またフレーズを聴く際も音程で把握しており、フレーズ再現も同様に行う。
- コード構成音はサクスの指の動きとして獲得している。
- フレーズの大まかな動き・形状はコードトーンやテンションを中心に「即興的に自分で定め」(周りの演奏の影響を大きく受ける)、その間をつなぐ細かい動きは「指に任せる」(周りの演奏には影響を受けない)。
- アウト感などは、フレーズを覚えて使うことで獲得してきた。
- 和音は響きを「感覚的」に捉えている。自分で楽器を(単音アルペジオで)鳴らすことで響きを習得してきたらしい。

### 3.3. ジャズギター学習者 (第三者としての聞き取り)

音楽の専門訓練は受けておらず、これまではロックギターを演奏していた。

- ギター指板上に音高がマッピングされた状態で音高を把握している。つまり単音スケールの指の動きと指板を対応させて音高を把握している。基音を元にしたスケールで音高を捉える傾向が強く、かつ認知が指板に依存する傾向があったため固定ドとも言える。
- 和声的短音階、旋律的短音階に関しては「スケール」として認知しており、自然にシャープするという感覚ではない。
- 固定ド、移動ドなどの階名でフレーズを歌うことができない。(歌詞があると歌うことができる)
- 和音の構成音という感覚がロックギターを演奏している時にはあまり意識しないでも演奏できたため、コード構成音に関しては、まだそれほど体得できていない。

### 3.4. ジャズギタープロ演奏家 (第三者としての聞き取り)

ギター以外の楽器経験やクラシック的な音楽情報の知識はない。国内のジャズギターのトップコンクールの優勝者。

- 音高はたとえ単音フレーズあっても「コードを抑える指の形がまず頭に浮かび、そこからどの指をどのぐらい動かした場所」という形で認識してい



図 1: 高音域に行きたい場合の、9度を中心とした旋律生成について、基音が一番下にはない押さえ方を元に説明するジャズギタープロ演奏家

る。単音フレーズ演奏の場合も「まず頭に指の形が浮かぶ」。

- コードトーンの基音抜きは容易にできる。基音を押さえている指を基音ありの指の形から抜くのではなく、基音抜きの指の形として覚えている。
- 移動ド傾向が強い。固定ドはギター指板があることにより成り立っている。
- 固定ド、移動ドなどの階名でフレーズを歌うことができない。
- アウト感も同様に「和音&指の形」として認識している。

## 4. 考察

記録から推測されるこれらの個々の演奏家の認知は、例えば両ギタリストの認知のようにギターのように楽器の特性に依存し楽器で共通している部分もあるものの、同じ楽器の演奏家でも、フレーズの認知ひいては音高の認知の形態が大きく異なっていることを示している。前述した著者のアマチュアベーシストを対象とした聞き取り調査の場合の場合ではアマチュアベーシストの多くが同じ回答をしていたのと、現在進めている調査では異なり、仮説である「フレーズひいては音高の認知は、楽器と身体の物理的特性に依存している」のみでは不十分であることがわかった。また、このような楽器以外の個人に依存する多種多様なフレーズひいては音高の認知のあり方が個人ごとにありうるため、楽器単位の調査では不十分であり、演奏者ごとに調査を行うことがあることを示している。

著者は、このことに注目し、このようなフレーズひいては音高の認知が「演奏の個性」を生み出しているのではないかと推測している。モデル化を行うため引き続き調査を行なっていく。

例えばジャズギタープロ演奏家は、音高の認識がコードを抑える指の形であることから、指板と指の身体的特性だけではなく、旋律の認知がコードを押さえる指のフォームにも強く影響を受けている。例えば一つの

コードを抑える場合、ルートなし、5th なしなども含めて複数の押さえ方があるが、それでも両手を使って鍵盤でコードを抑えるのに比べ自由度が低い。

しかしながら、このジャズギタープロ演奏家のフレーズ・音高の認知は極めて特殊な形をしているからこそ、フレーズ生成モデルに「自然な偏り」を持たせることができる、と著者は考えている。

またジャズサクソプロ演奏家の場合では、高速に指が動かす場所はほぼ自動的に行われている、つまり身体知として獲得されている。しかしながらフレーズ全体が自動生成であるとはつまらない演奏であるとし、フレーズの形状は自分でその場で、周囲のプレーヤの演奏を聴きながら決めることを重要視している。

自らその場で即興で先導するフレーズの大まかな形状を生成する箇所と、「指に任せる」すなわちアルペジオやスケールの行き来を乱数的に生成する箇所と箇所は、別のルールとして定めるべきであり、この2つを行き来することが、生成モデルに「自然な偏り」を持たせることができると著者は考えている。

ルールベースと乱数生成を組み合わせる旋律生成モデルを作ろうとする場合、ルールの縛りが自由であればあるほど乱数に依存する傾向が強くなり、不自然さが強くなる。これを「新たな創作可能性」として見ることも作曲家の立場からは行われているが、しかし演奏家から見れば演奏しづらい、演奏表現を行いつづらぬ旋律が生成されてしまう。できるだけ演奏家が自然な演奏表現を行える範囲内で「新たな創作可能性」を追求することが、少なくともジャズの旋律生成では重要であると著者は考える。

また、作曲家が考えた生成ルールでは五線譜表現になってしまい、楽器の物理的な制約などは禁則事項としてルールに取り込まれるだけのことが多い。実際にはジャズサクソプロ演奏家もジャズギタープロ演奏家も、「本人が獲得している指の形」をモデルに取り込むべきであり、楽器の制約をルールに取り込むだけでは不十分だと著者は考える。

## 5. まとめ

著者は、ジャズ即興演奏人工知能と人間の演奏家が「音楽的知識を共有すること」を重要視し、人間の演奏家の演奏生成モデル化のための調査を行っており、本稿ではその第1報としてフレーズひいては音高の認知について、現段階の報告を行なった。

生成モデルは楽器の物理的制約のみに依存し、同じ楽器の演奏家・学習者であれば同じような傾向があると予想していたが、実際には同じ楽器の演奏家・学習者であっても大きく傾向が異なることがわかった。身体知として獲得している「指の形」が、何らかの形で

フレーズ・音高の認知に大きな影響を及ぼしていると著者は推測している。

今後は、サクソについては一人名記述および二人名聞き取り調査を進め、ギターは第三者的な観点から被験者を増やして、同じ演奏家に師事した学習者同士でも認知やその形成が異なるかどうかを検証していく。

## 6. 参考文献

Esling, P., A. Chemla-Romeu-Santos, and A. Bitton (2018). Generative timbre spaces: regularizing variational auto-encoders with perceptual metrics. <https://arxiv.org/abs/1805.08501>.

Sudnow, D. (1978). *Ways of the hand: the organization of improvised conduct*. Cambridge, Massachusetts: Harvard University Press.

安藤大地 (2015). ジャズ即興演奏における4ビートベースラインの2段階対話型遺伝的プログラミングによる生成. In 情報処理学会インタラクシオン2015 シンポジウム.

## 7. 著者プロフィール

### 安藤 大地 (Daichi ANDO)

国立音楽大学声楽科を経て音楽デザイン学科卒業。スウェーデン Chalmers University of Tehcnology より MSc. 授与。東京大学大学院新領域創成科学研究科基盤情報学専攻博士課程終了。博士(科学)。現在首都大学東京システムデザイン学部インダストリアルアート学科助教。

声楽を松本進に、作曲とコンピュータ音楽を Palle Dahlstedt と葉孝之、Cort Lippe に、クラシックサクソフォン穴戸陽子に、ジャズサクソフォンを尾崎朝子にそれぞれ師事。

対話型進化システムや人工生命、群知能などの音楽創作分析への応用を中心とした研究を行っている。また自らの研究成果を応用した音楽作品、メディアアート作品の創作を行っている。



この作品は、クリエイティブ・コモンズの表示 - 非営利 - 改変禁止 4.0 国際 ライセンスで提供されています。ライセンスの写しをご覧になるには、<http://creativecommons.org/licenses/by-nc-nd/4.0/> をご覧頂くか、Creative Commons, PO Box 1866, Mountain View, CA 94042, USA までお手紙をお送りください。