

Linguitone: 音色の言語表現を用いた エレクトリックギターの音作り支援システム

小林 篤矢^{1,a)} 徳井 直生^{2,b)}

概要: エレクトリックギターの演奏において重要なスキルの一つであるエフェクトペダル等を用いて音色を変化させる「音作り」には、音色の特徴を表すのに独特な言語表現が用いられる。そのため、初心者は音色の言語表現と実際の音像が結び付かず、具体的な音をイメージしエフェクターを用いて音を作ることが難しい。本研究では、音色の言語表現から音作りを支援するシステム **Linguitone** を提案する。当システムの開発にあたり、ギタリストへの事前アンケート調査、ギターの音色の言語表現のデータの収集、データセット作成用インターフェースの作成、自然文の分散表現からエフェクトのパラメータを予測する機械学習モデルの構築を行い、Web ブラウザで動作するインターフェースを開発した。さらに、開発したシステムに対してはギター奏者を被験者としたユーザビリティと音作りの支援性、創造支援性の評価を実施した。

Linguitone: A Web-based System to Assist in Making Electric Guitar Sounds Using Natural Language Expressions

ATSUYA KOBAYASHI^{1,a)} NAO TOKUI^{2,b)}

1. はじめに

エレクトリックギターは基本的にアンプに接続して音を出す。その途中に、エフェクトペダル（エフェクター）と呼ばれる機器を挟むことによってそのサウンドに変化をつけることができる。ギタリストが出したい音色を求めてギター本体やアンプ等の音響機材の調整を行う「音作り」と呼ばれる作業は、演奏技術の一つともされている。調律師や個体によって基本的な音色が定まるピアノ等とは異なり、エレクトリックギターの場合はこの機材による音作りの幅が非常に広く、音作りのノウハウをまとめた書籍や著名ギタリストの機材とその使い方のみに焦点を当てた雑誌なども多数存在する。

その音作りによって得られるサウンドの特徴を表す際には、独特な言語表現が用いられている。例えば、代表的なエフェクターブランドであるローランドの BOSS の製品紹介^{*1}には以下のような表現がある。

製品紹介のフレーズ

『OD-3 の周波数特性は幅広く、太く芯のある低域と美しい高域、粘りのあるサステインを実現』
『豊かなゲインとサステインが得られます』
『ブライトでハリのあるカッティング・サウンドやファットで迫力ある低音弦リフ、伸びやかでツヤのあるリード・サウンド』

エレクトリックギターを長年演奏していない人からすると、これらの表現を目にして明確な音色をイメージすることは難しいと考えられる。そこで、ここまで述べたエレクトリックギターの音色に関する言語表現（音色表現句）について、次の 2 つの仮説を立てた。

- 仮説 A：エレクトリックギターの音色表現句は実際

¹ 慶應義塾大学 環境情報学部
Faculty of Environment and Information Studies, Keio University

² 慶應義塾大学大学院 政策・メディア研究科
Computational Creativity Laboratory, Graduate School of Media and Governance, Keio University

a) t16366ak@sfc.keio.ac.jp

b) tokui@sfc.keio.ac.jp

^{*1} <https://www.boss.info/jp/>

の音をイメージできるようになるまでに長年の経験が必要であり、初心者にとっての音作りの困難さにつながっているのではないかと

- 仮説 B: 仮説 A が正しい場合、音色表現句からエフェクターの設定を行い音作りを自動化するシステムによって、エレクトリックギター初心者の練習・演奏の支援を行うことが可能なのではないかと

本研究では、上記の両仮説の正当性を示す上で、エレクトリックギター奏者への調査とシステムの実装と評価実験まで行う。

2. 関連研究

2.1 音・音色・音楽を表す表現語について

音や音楽について説明・表現して他者へ伝達するために、言語による表現を行う例は多くある。曾根ら（1962）は、一般的な語と音楽関係や音響関係から表現語を集め、その中から選択した高い使用頻度が確認された 21 の表現語を用いて 7 段階の音源聞き取り評価を、10 秒程度の 8 種類の音源に対して行い因子分析を行っており [1]、宮原ら（1996）は、オーディオ雑誌内の記述から音質に関する評価語を調査し、192 個の評価語を抽出し、それらの各群の代表語に対して様々な音質の音楽 CD の聴取評価実験を行い分析を行っている [2]。楽器の音では、ピアノの音色表現語とそれに紐づく奏法について分析している後藤の研究（2018）[3]があり、ピアノの調律は専門家が行うためその時点で音作りは完了しているのにも関わらず明確に奏法によって出てくる音色が異なって認知されレッスンの場などでも音色表現語を用いた指導がなされている現状に着目している。ピアノを指導しているピアニスト良い音色と良くない音色の双方についての表現語について品詞を用いた分類を行っている。

さらに、音楽の特徴や聴取時の感性を表す言語表現の分析やそれらを活用したインターフェースやシステムも提案されている。河村ら（2003）は音楽雑誌 rock'in on JAPAN などから 530 語の感性語を抽出し、それらの語から音楽がイメージ可能かどうかをアンケート調査の実施後に、先行研究では取り上げられていない感性語を洗い出した [4]。これらの音楽聴取時の感性語に関する研究は、感性語をクエリとする音楽データベース [5] や楽曲検索システム [6] の研究に応用されており、感性語を用いた SD (Semantic Differential) 法により空間を構築し、その感性空間内での演算により似た楽曲を選出している。感性語を入力とすることで、ユーザーが対象となる楽曲の特徴の詳細を知らなくても感覚的に探すことを可能にしている。

2.2 エレクトリックギターの演奏支援

エレクトリックギターの音色と言語表現の関係性については Marui と Martens（2005）の研究 [7] が挙げられる。

エレクトリックギター用エフェクター^{*2}を用いて歪みの強さ (Gain) と音色の明るさ (Sharpness) がそれぞれ 3 段階ずつ計 9 種のエフェクトを利用し、それらの音刺激に対して 9 つの評価語対 (表 1) による評定を 57 人の被験者に行わせている。

表 1 Marui と Martens（2005）が用いたエレクトリックギターのディストーションサウンドへの評価語対

弱い	強い
粗い	きめの細かい
広がりがある	やせた
快い	不快だ
軽い	思い
かたい	やわらかい
澄んだ	濁った
騒々しい	静かな
薄い	厚い

この調査結果に個人差多次元尺度構成法 (INDSCAL) を適用した 2 次元へのマッピングによると、「弱い～強い」が最も音の印象の差と相関があったため、音の明るさと音の強さの 2 軸によってエフェクトのパラメータを操作するより直感的なインターフェースが実現可能であると指摘している。

また、ソフトウェアによる演奏スキル習得に関する研究としては、元川ら（2007）の 3DCG による押弦パターンを Web カメラで撮影中のユーザーのギターの指板上にリアルタイムで表示することで、ユーザーは 3DCG の手に自身の手を重ね合わせるようにして押弦の練習ができるシステム [8] が挙げられる。Web カメラで撮影した映像に対して、マーカーとギターの指板から位置の推定を行い、正確なコード名と 3DCG の手の表示を可能にしている。評価実験として 8 名のギター未経験者を被験者に、他の練習システムとの比較を行っている。9 種類のコードに対して、正しくコードを演奏できるようになるまでの時間と、ディスプレイとギターの間の視線移動の回数を計測し、使用感のアンケートも取得している。

また、Web アプリケーションという形をとったものでは有賀ら（2017）が提案する Strummer[9]があり、データから取得した各コードの出現率と難易度から重要性を示す Chord Primariness を算出することで、どの楽曲を次に演奏すべきかが判りやすい。インターフェースは Web アプリケーションとなっており、ユーザーの演奏に対してインタラクティブにフィードバックを行うことができる。ユーザーテストとして、5 名のギター演奏歴が無い被験者に対してそれぞれ 5 日間に渡りそれぞれ計 5 時間の練習時間を課し、課題曲の演奏を行わせている。

^{*2} MIDI によってコントロールができるラックマウント式エフェクター BOSS GX-700 を利用

しかし、これらの音の表現語に関する研究で扱われている語はどれも、先述の BOSS OD-3 の製品紹介にあるような独特の用語や表現ではなかった。

3. 事前調査

仮説 A を示すため、10 人のエレクトリックギター演奏経験者（経験年数 半年～35 年、平均 8.0 ± 9.35 年、中央値 5.0 年）を対象に、エレクトリックギターのサウンドに特有な言語表現についてのアンケート調査（表 2）を実施した。調査はオンライン上で実施し、エレクトリックギターに特有の表現とそれらの表現語を実際のサウンドと紐付けて理解できるようになるまでにかかった時間の回答を得た。

表 2 ギター奏者へのアンケート項目

1	エレキギター演奏歴を教えてください
2	あなたの思う「エレキギターのサウンドに特有な言語表現」を教えてください。
3	以下の言語表現のうち、具体的なサウンドがイメージできるもの音作りのイメージができるものにチェックを付けてください。
4	以上のような言語表現をイメージできるようになるまでギターを始めてから大体どのくらいかかりましたか？

3.1 エレクトリックギターに特有な言語表現について

各回答者から以下の音色表現句が「エレクトリックギターに特有な表現として思い浮かぶもの」として挙げられた。

- ・ ジューシーな歪み
- ・ ギュイン
- ・ 粒立ちのよい歪み
- ・ ジャキジャキ、ベケベケ
- ・ うねりのあるサウンド
- ・ 潤いのある歪み
- ・ 芯のあるウォームサウンド、ブリティッシュサウンド
- ・ エッジの効いたカッティングリフ
- ・ ギターの割に音が大きい。ダイナミクスレンジが広い。
- ・ ピッキングの強弱から見られる有機性

3.2 音色表現句からの具体的な音色のイメージ

楽曲の感性語に関する研究 [4][6] での感性語の評価手法を参考に、具体的な音色のイメージが可能かどうかを調査した。代表的なエフェクターブランド各社（BOSS^{*3}、MXR^{*4}、Ibanez^{*5}、Fender^{*6}）の製品紹介ページから以下の

^{*3} BOSS Stompboxes: コンパクト／ツインペダル <https://www.boss.info/jp/categories/stompboxes/>

^{*4} モリダイラ楽器 Effectors for Guitar, <https://moridaira.jp/mxr/for-guitar-mxr>

^{*5} ELECTRONICS — PRODUCTS — Ibanez guitars, <https://www.ibanez.com/jp/products/category/electronics/>

^{*6} Effects Pedals — Fender, <https://shop.fender.com/ja-JP/effects-pedals/>

言語表現（表 3）を手動で取得し、これらの表現に対して「具体的なギターサウンドのイメージが可能か」と、「音作りのイメージができる」かどうかの聞き取りを実施した。

表 3 各社エフェクターペダルの製品ページにある音色表現句

音色表現句	メーカー
豊かな倍音	BOSS
ダイナミックなサウンド	BOSS
滑らかなサステイン	BOSS
心地よくマイルドな揺れ	BOSS
美しく広がる	BOSS
深い揺れ感	BOSS
変調感の強いくせのあるサウンド	BOSS
爽やかな	BOSS
立体感のあるステレオ	BOSS
厚み	BOSS
ウォームで粘りある	Ibanez
ざらついた歪み	Ibanez
ナチュラルで温かみのあるトレモロ	Ibanez
ファットで暖かみのあるコーラス	Ibanez
伝説的な焼けつくようなトーン	MXR
明瞭で煌びやかなサウンド	MXR
クリアかつワイドレンジな解像度の高いサウンド	MXR
ジェット音やスペースエフェクト	MXR
リッチかつクリーミーなサウンド	MXR
分厚く迫力に満ちた	Fender
分厚く強烈なサチュレーションの効いた	Fender
スムーズで上品なサウンド	Fender
サウンドに独自のスピンを作り出し	Fender

回答結果として、図 1 が得られた。今回の回答者 10 人では、4 項目を除き過半数が「サウンドがイメージできる」状態にあったが、ほとんどの表現句で「音作りのイメージ」が可能な回答者は半数程度に留まった。

3.3 音色表現句と具体的な音像との結びつきの獲得

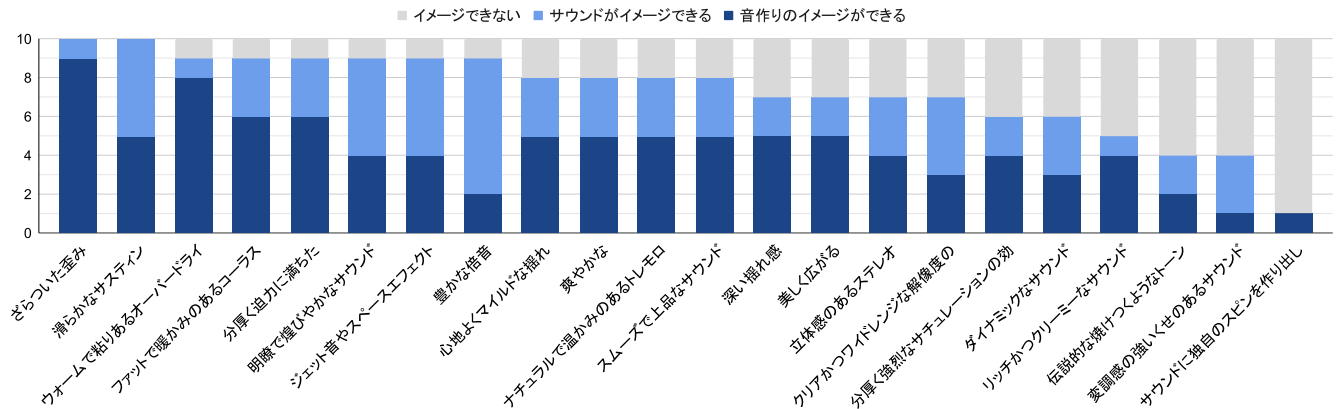
「言語表現をイメージできるようになるまでにかかった時間」の調査結果としては、以下の回答が得られた。

- ・ 10 名のうち 2 名（それぞれ経験年数半年と 4 年の回答者^{*7}）が「未だにわからない」と回答した
- ・ 1 名から「表現のイメージはできるが自身での表現はしない」という回答が得られた。
- ・ 1 名から「最近わかりはじめた」という回答が得られた。この回答については、経験年数を「イメージできるようになるまでにかかった時間」として扱った。

上記の 7 名からの回答のデータから平均で 4.0 年かかるとのデータが得られた。正規分布に従うと仮定すると、95%信頼区間は 2.93 年から 5.07 年の間であり、最低でも

^{*7} この 2 名は「音色表現句に対してのギターサウンドのイメージの可否」に対して「イメージできる」と回答している項目があり、質問の意図を理解していないか自身の音色表現句の理解度への自信が無いものと思われる。

図 1 ギター奏者へのアンケート調査：音色表現句に対してのイメージ可/不可の回答



3 年の演奏経験が必要であると解釈できる。

以上の事前調査から、エレクトリックギターの音色の言語表現には特異な言語表現（音色表現句）が見られ、それらの音像との結びつきを得るには演奏経験としてある程度多くの時間が必要であることが判明した。またこの結果は、音色表現句から具体的な音作りのプロセスをイメージできる奏者が約半数程度に留まる可能性を示唆しており、初心者でないギタリストにとっての音作り支援の必要性を裏付けるものであると考えられる。

4. Linguitone の実装

4.1 各モジュール

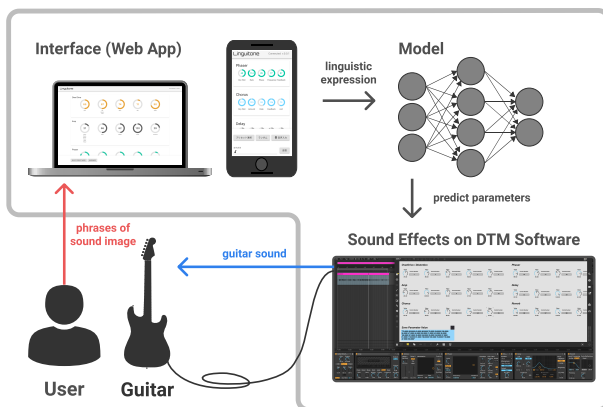


図 3 システム概要図

主に 3 つのモジュールに分けて、提案システムの実装を行う。エレクトリックギターからの入力オーディオインターフェースを通し DTM ソフトウェアである Ableton Live^{*8}へと入力される。Ableton Live に同梱されているエレクトリックギター向けオーディオエフェクトとアンプシミュレーターを用いて音作りを行う。このエフェクトの各パラメータの操作を、音響プログラミング環境である Cycling '74 Max^{*9}のパッチにより自動化する。また、この

自動化のためのパッチはデータセットの作成にも用いる。

ユーザーからの操作を受け付けるインターフェースは、GUI を設計する上での効率と拡張性・公開のし易さから Web アプリケーションの形を取る。

各エフェクトのパラメータを予測する機械学習モデルと Web アプリケーションのインターフェースとの通信を行うサーバ部、音響プログラミング環境へのパラメータの送信部は Python で実装した。

4.2 ソフトウェアによるギターエフェクトシステム

7 種類の Ableton Live のオーディオエフェクトをと、パラメータ固定のキャビネットシミュレーターを利用した。それぞれのエフェクトのどのパラメータを学習するかは、エレクトリックギター向けエフェクトペダルで主流となっている操作パラメータをそれぞれ 1 種類につき 5 つずつ選択した。

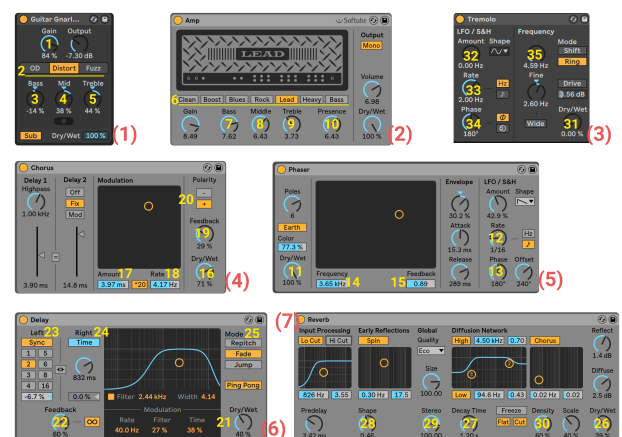


図 4 使用する Ableton Live Suite の audio effect. (1)Pedal, (2)Amp, (3)Tremolo, (4)Chorus, (5)Phaser, (6)Delay, (7)Reverb に対して 5 つずつ全 35 パラメータを割り振り、それらを操作し音作りを行う。

^{*8} Ableton Live <https://www.ableton.com/en/live/>

^{*9} Cycling '74 Max <https://cycling74.com/products/max>



図 2 Linguitone のユーザーインターフェース (スマートフォン)

5. データセットの作成

5.1 音色表現句の抽出

学習データ作成のために、エレクトリックギターに特有な音色表現句の抽出を行いたい。そこで、エフェクターの商品レビュー文章に着目し、大手楽器通販サイトであるサウンドハウス^{*10}と、デジマート^{*11}から、1519 種類のギターエフェクターに対する 5187 件のカスタマーレビューのテキストデータを取得した。

収集したレビューには、音色に関係の無い価格や配送についての文章も含まれる。そこで、小林ら (2005) による Web 上のデータからの評価表現の自動抽出 [10] での品詞によるフィルタリング手法と [11] [12] で提案されている係り受け解析によるキーワード抽出手法を参考に、事前に抽出した音色表現句 (表 3) の品詞列のパターンと、「音色」「響き」「サウンド」「エフェクト」や各エフェクターの名前等のキーワードに対して係り受けている文章を抽出した。係り受け解析には日本語自然言語処理ライブラリである Sudachi[13] と Ginza^{*12}を用いた。

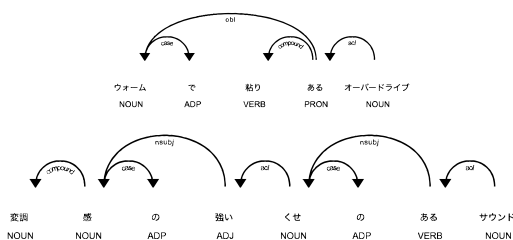


図 5 音色表現句への係り受け解析の結果の例。このようにエフェクト名や「サウンド」などのキーワードに係るフレーズを抽出した。

レビューデータの全 33,623 文の中から、係り受け構造を

^{*10} サウンドハウス - PA 音響機器, DJ・DTM, 照明機器, ステージ, カラオケ機材の全国通販
<https://soundhouse.co.jp>

^{*11} 【デジマート】楽器と音楽機材専門の検索&総合情報サイト
<https://www.digimart.net>

^{*12} GiNZA - Japanese NLP Library - <https://megagonlabs.github.io/ginza/>

もとにした抽出で 3,466 件を抽出し、さらにその中から上記の品詞列パターンを用いて 899 件の文に絞り込みを行った。そこから、エレクトリックギターの音色表現句であるフレーズを手作業で抜き出し、最終的に 170 件のフレーズを抽出した。そこへ事前に取得していた 23 件 (表 3) と、アンケートから得られた 10 件を追加した後に重複を削除し、198 件の音色表現句を抽出した。

5.2 エフェクトパラメータデータの作成

収集した音色表現句に対応する音色を再現し、その時の各エフェクターのパラメータを保存することでデータセットを作成する。Cycling '74 Max によるデータ作成用インターフェースを用いて音作りを行った。

加えて、上記のみの手法ではデータ作成にあまりに時間がかかるうえに著者による音色表現句の感覚のみが反映されてしまうため、音色表現句 → エフェクトパラメータという方向のデータ作成だけでなく、エフェクトパラメータ → 音色表現句という方向でのデータ作成を行った。筆者が作成したエフェクトパラメータがそれぞれ正規分布に従うと仮定しサンプリングを行い、得られたエフェクトパラメータによるギターの音色を聞き取り、音色表現句のアノテーションを行うことでデータ作成を行った。アノテーションはエレクトリックギター演奏経験のある成人男性 2 名 (平均経験年数 7.75 ± 2.3) に依頼し、追加で 100 件のデータの作成を行った。

さらに、作成したデータに対して簡単なデータ拡張として、音色表現句のデータに対して音色に影響しない同義語 (表 4) の挿入を行った。最終的に、100 件のデータ追加を行い、最終的に 311 件のペアデータを取得した。

6. モデリング

6.1 言語モデル

システムで入力されるのは音色表現句のテキストであり、モデルも自然言語での入力を想定して作成する。自然言語を定量的に扱うために、単語の埋め込み (分散表現) を用いる。

表 4 データ拡張に用いた同義語群

サウンド	音	
音色	トーン	
変調	モジュレーション	
歪み	ドライブ	
太い	ファットな	
位相	フェイズ	
ウォーム	温かい	あたたかみ
尖った	シャープ	するどい

単語の埋め込み手法の際に着目した研究 [14] を参考に、分散表現手法として tf-idf, fastText, Doc2Vec の比較を行う。文章 (n 個の単語の系列) S を受け取って、次元数 $D \in \mathbb{N}$ の単語ベクトル $\mathbf{x} \in \mathbb{R}^D$ の系列 ($N \times D$ の行列 $\mathbf{X}$) を返すモデル $E_{\text{sequence}}: S \mapsto \mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$ もしくは、単語ベクトルと同一次元のパラグラフベクトルを返すモデル $E_{\text{paragraph}}: S \mapsto \mathbf{x}$ を実装し。形態素解析には MeCab[15] を用い、辞書には IPA 辞書と NEologd 辞書 [16][17] を、学習済み単語埋め込みモデルに用いられているものに合わせて適用した。

6.2 エフェクトパラメータ予測モデル

今回扱った 7 種類のエフェクターの計 35 次元のベクトル $\mathbf{p} \in \{0, \dots, 127\}^{35}$ として扱い予測モデルを作成し、複数の手法から最適なものを選択した。多重出力の L1, L2 正則化付き線形回帰 (Elastic Net) と 3 層ニューラルネットワーク, LSTM 層を用いたニューラルネットワークによる非線形回帰をそれぞれ実装した。モデルの選択基準には、予測するベクトルの次元毎に重みをつけた Weighted Mean Squared Error (wmse) を用いる。この重み $\mathbf{w}^{\text{Loss}}$ は、予測の精度が異なると作られるサウンドに大きく影響を及ぼすと考えられる各種エフェクトの Dry/Wet を操作する次元を 2.0 とし、その他は 1.0 とした。

$$\mathbf{w}^{\text{Loss}} \in (0, 2]^{35} \quad (1)$$

$$\text{wmse} = \frac{\sum_{k=1}^{35} \mathbf{w}_k^{\text{Loss}} (\mathbf{p}_k - \hat{\mathbf{p}}_k)^2}{\sum_{k=1}^{35} \mathbf{w}_k^{\text{Loss}}} \quad (2)$$

表 5 各分散表現手法と各機械学習モデルの精度 (wmse)

	Elastic Net	3 層 NN	LSTM
tf-idf	459.04	1121.70	N/A
fastText	460.20	354.65	461.49
Doc2Vec	455.69	480.38	N/A

結果 (表 5) として、fastText によるベクトル化とニューラルネットワークの組み合わせが最も精度が高かった。よって、このモデルを Linguitone へ採用し、ユーザーインターフェースモジュールからの入力に応じて推論が行われるよう実装

した。

7. エフェクト

Linguitone 利用時のエフェクトのかかり具合を、440Hz の単音に対する波形の変化を見て比較する。入力された音色表現句の違いで、歪みのゲイン量が増減し、倍音成分の振幅が大きく変化する様子が分かる。

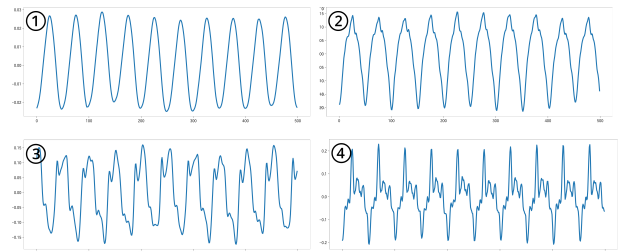


図 6 音色表現句と出力音の波形: (1) エフェクト無し, (2) 「分厚いサチュレーションのオーバードライブ」, (3) 「ざらついた歪み」, (4) 「メタリックなサウンド」

8. 評価

8.1 被験者と実験環境

ギターの演奏経験が 4 年以上ある成人男性 6 名 (平均経験年数 7.0 ± 2.24) から、当システムの利用後の評価質問紙への回答を得た。被験者には簡単な操作説明を行った後にモニタースピーカーの前に座ってもらい、用意された実験用 PC に接続されたエレクトリックギターを演奏しながら Linguitone を PC とスマートフォンから操作し、一人あたり 15 分程度自由に音作りと演奏を実施してもらった。エレクトリックギターは、Fender Japan Stratocaster と、Fender USA Telecaster Custom を利用した。

実験後に、ユーザビリティ評価、音作りの支援性、創造支援性の 3 つの評価質問紙に対しての 5 段階のリッカート尺度を用いた評価を行ってもらい、回答結果に以下の点数を付与した。

- そう思う: 5
- どちらかといえばそう思う: 4
- どちらとも言えない: 3
- どちらかといえばそう思わない: 2
- そう思わない: 1

また、被験者 5 と被験者 6 に対しては、インタビューによるフィードバックを得た。

8.2 ユーザビリティ評価

Linguitone のソフトウェアのユーザビリティ評価には、ユーザー体験 (UX) に基づいてインタラクティブシステムを評価する質問紙 [18] で提案されている評価項を用いた。結果は以下のようになった。

- (1) 理解しやすい 平均 4.50 ± 0.84

- (2) 容易に修得（習得）できる 平均 4.00 ± 1.10
- (3) 迅速に応答する 平均 4.83 ± 0.41
- (4) 期待した通りに動かせる 平均 3.17 ± 1.33
「期待した通りに動かせる」のみ 4.0（どちらかといえば
そう思う）を下回った。

8.3 音作りの支援性評価

システムによる音作りの支援性の評価のために、以下の評価項目を設定した。結果は以下ようになった。

- (1) 音作りのイメージを容易にする 平均 4.00 ± 1.10
- (2) 音作りのスピードを上げる 平均 3.67 ± 1.21
- (3) 音作りの幅を広げる 平均 4.00 ± 1.26
- (4) 音作りに関する知識が無い初心者のサポートとなる 平均 4.50 ± 1.22
- (5) 自身がギターを始めてすぐの頃にあったら使う 平均 3.67 ± 1.37
- (6) 今でも使いたい 平均 3.50 ± 1.05

本研究の仮説を示す上で重要な、「音作りに関する知識が無い初心者のサポートとなる」の得点が高く得られた。一方で、「音作りのスピードを上げる」への評価は高くなく、被験者が実際にユーザーとして使いたいが現れる「自身がギターを始めてすぐの頃にあったら使う」と「今でも使いたい」の得点の平均も 4.0 を超えていない。

8.4 創造支援性

このシステムの創造支援性について、創造行為を支援するシステムの評価を行うための指標として提案されている Creativity Support Index[19][20] から Linguitone の評価に適切な項目を選定した。結果は以下ようになった。

- (1) オプション、デザインを探っていくのは簡単 平均 3.66 ± 1.03
- (2) さまざまなアイデアや成果を追跡するのに役立つ 平均 4.16 ± 0.41
- (3) クリエイティブな活動ができる 平均 4.00 ± 1.10
- (4) 表現力を発揮することができる 平均 4.17 ± 0.75
- (5) 集中して使うことができる 平均 3.67 ± 1.03

「さまざまなアイデアや成果を追跡するのに役立つ」、
「表現力を発揮することができる」にて高い評価を得た。

9. 考察と展望

被験者への使用感のインタビューで得られたフィードバックから本システムの問題と改善点について考察を行う。

個別のエフェクターやアンプにおいて想像される音像、例えば「チューブスクリーマーの音」とか、このエフェクターは誰々が使ってる、とか、エフェクターやアンプにおいて求めることはそういうギアかつ精神的なもののインストールから始まっているような気がする。（被験者 5）

確かに、エレクトリックギターの初心者が音作りを始めるうえで代表的なエフェクターやアーティストに象徴される固有の音を出せるようになることも重要であるが、Linguitone が目指していることはサウンドのより細かなニュアンスのコントロールを、他のギタリストとのコミュニケーションの中で行えるようになることである。具体的に言えば、スタジオで練習している際に自身のギターの音色に対して「もう少し滑らかな歪みで、丸みの帯びたトーンがほしい」*13などのコメントがあった際に、自身の感覚で適切なパラメータを操作できるようになることである。また、固有名詞に象徴される音色も音色表現句の組み合わせによって表現可能ではあるが、それを Linguitone 上で実現するのは非常に難しいと考えられる。固有名詞間の意味的な違いを適切に表現することは難しく、「ジミ・ヘンドリックスのサウンド」という入力から直接ファズを使用した激しいサウンドを導出しながらも、「エリック・クラプトンのサウンド」という入力に対して的確な音色を表現する事は、現状のアルゴリズムでは困難であると考えられる。今回活用した IPadic NEologd による形態素解析では固有名詞を単語として扱うことができるが、Wikipedia で学習された分散表現ではジミ・ヘンドリックスとエリック・クラプトンは非常に似たベクトルを持っており、両者の差分を明確に取り込むのは難しい*14。

出音は自分のイメージに沿ったもので、大きくハズレているという印象は全くなかった。しかし、キャラクターが突き抜けておらずメリハリがない。それぞれのエフェクトがすべて微妙にかかっている。それだけのような音になることも多かったと感じる。（被験者 6）

また、被験者 6 によるコメントには、エレクトリックギター経験者の音作りに関する言語感覚をツールを通して初心者にも提供できることが示されている一方で、言語の表現から得られるエフェクトに大きな差が出にくいという意見もあった。これは、学習に用いたデータのサンプル数が不足モデリングの精度が低いことが原因として考えられる。現状のサンプリング手法では平均的な音となるパラメータ付近のデータが多くなるので、一様分布でサンプリングしアノテーションを行うなどのデータ作成をする等が効果的であると考えられる。

また、本研究ではエフェクターのみの調整を扱っているが、ギター自体の音響特性やスピーカーキャビネットの調整などの他にも多くの要素がギターの最終的な音色に影響する。そのため、今後はギターのピックアップの種類や位置の推定やキャビネットの種類など他のパラメータを予測

*13 筆者が実際に言われた経験がある

*14 音楽雑誌上の文章によるコーパスを作成し、そこから Skip-Gram 等でモデルを作成すれば、音楽ジャンルや年代等の表現が得られる可能性がある

するためのより細かいデータ収集が必要となる。

アクセシビリティの観点では、少なくとも現状の Linguitone は、Python の環境・Max・Ableton Live Suite が揃っており、かつ日本語自然言語処理のための数 GB の学習済みモデルファイルやその他多数のライブラリをインストールしていないと動作させることすら出来ない。AmpliTube^{*15}や Positive Grid Bias^{*16}等の一部が無料で利用できるソフトウェア演奏環境を想定したシステムを開発すれば、MIDI の設定項目を公開した上で Web MIDI API を組み合わせれば、Linguitone を Web 上にホスティングして公開することが可能であった。

10. おわりに

本研究では、エレクトリックギターの音色表現句に着目し、ギタリストへのアンケート調査の後に係り受け解析と品詞パターンマッチによるエフェクターの商品レビューデータからのフレーズ抽出を行い、そこからエフェクトパラメータのペアデータセットの作成とそれによる予測モデルの構築、Web アプリケーションベースのインターフェースの実装を行うことで、自然言語による音色の表現からエレクトリックギターの音作りを行うシステムである Linguitone を開発した。

さらに、Linguitone のねらいは、3 年以上の演奏経験が必要と考えられる音色表現句からの実際のエレクトリックギターの音色のイメージを行う能力がまだ無いエレクトリックギター初心者・中級者のための音作り支援であることから、アプリケーションのユーザビリティ評価に加え支援の有効性と創造支援性の評価を行った。

また、作成したデータセットからエレクトリックギターの各種エフェクトに対して関連のある音色表現句の分析を行い、システムの有効性とソフトウェアによるエレクトリックギターのエフェクター実装の課題と Linguitone のソフトウェアとしての可用性について考察を行った。

参考文献

- [1] 曽根敏夫, 城戸健一, 二村忠元: 音の評価に使われることばの分析, 日本音響学会誌, Vol. 18, No. 6, pp. 320–326 (1962).
- [2] 宮原 誠, 守田幸徳: 音質を表現する評価語の調査分析, 日本音響学会誌, Vol. 52, No. 7, pp. 516–522 (1996).
- [3] 後藤友香理: ピアノの音色表現語に関する研究—ピアノ・レッスンに着目して—, 『教育学部研究報告 (人文・社会・自然科学篇)』第 68 号, Vol. 68, pp. 185–196 (2018).
- [4] 河村知典, 杉原太郎, 森本一成, 黒川隆夫: 733 音楽から受ける印象の評価語に関する検討, *Dynamics & Design Conference*, Vol. 2003, pp. 733-1–733-5 (2003).
- [5] 池添 剛, 梶川嘉延, 野村康雄, Others: 音楽感性空間を用いた感性語による音楽データベース検索システム, 情報

- 処理学会論文誌, Vol. 42, No. 12, pp. 3201–3212 (2001).
- [6] 杉原太郎, 森本一成, 黒川隆夫: m-RIK: 個人の感性特性に対応可能な音楽検索システム, 情報処理学会論文誌, Vol. 46, No. 7, pp. 1560–1570 (2005).
- [7] Marui, A. and Martens, W. L.: Timbre of nonlinear distortion effects: Perceptual attributes beyond sharpness, *Proceedings of the 2005 Conference on Interdisciplinary Musicology (CIM05)* (2005).
- [8] 元川洋一, 斎藤英雄: 拡張現実表示技術を用いたギターの演奏支援システム, 映像情報メディア学会誌, Vol. 61, No. 6, pp. 789–796 (2007).
- [9] Ariga, S., Goto, M. and Yatani, K.: Strummer: An interactive guitar chord practice system, *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1057–1062 (2017).
- [10] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: 意見抽出のための評価表現の収集, 自然言語処理, Vol. 12, No. 3, pp. 203–222 (2005).
- [11] 河野一志, 町田 翔, 村松拓実, 小林美穂, 延澤志保: 不特定分野の商品レビューを対象とした評価情報の自動認識, 第 79 回全国大会講演論文集, Vol. 2017, No. 1, pp. 535–536 (2017).
- [12] 木村優介, 楠 和馬, 馬場睦也, 波多野賢治: 係り受け解析を用いたドメイン固有のキーワード抽出手法の提案, 情報・システムソサイエティ特別企画学生ポスターセッション予稿集, 電子情報通信学会, p. 207 (2020).
- [13] Takaoka, K., Hisamoto, S., Kawahara, N., Sakamoto, M., Uchida, Y. and Matsumoto, Y.: Sudachi: a Japanese Tokenizer for Business, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (chair), N. C. C., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Hasida, K., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odiijk, J., Piperidis, S. and Tokunaga, T., eds.), Paris, France, European Language Resources Association (ELRA) (2018).
- [14] 佐藤一輝, 尾崎知伸: 単語埋め込み技術の違いによる日本語ツイートの感情分類精度の比較実験, 人工知能学会全国大会論文集, Vol. JSAI2017, pp. 1K11–1K11 (2017).
- [15] KUDO, T.: MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.sourceforge.net/>, (online), available from <https://ci.nii.ac.jp/naid/10019716933/> (2005).
- [16] Toshinori, S.: Neologism dictionary based on the language resources on the Web for Mecab (2015).
- [17] 佐藤敏紀, 橋本泰一, 奥村 学: 単語分かち書き用辞書生成システム NEologd の運用—文書分類を例にして—, 研究報告自然言語処理 (NL), Vol. 2016-NL-229, No. 15, pp. 1–14 (2016).
- [18] Matsumoto, K. and Zempo, H.: Questionnaire for evaluation of interactive systems based on user experience, *Jpn. J. Ergon.*, Vol. 53, No. 2, pp. 46–50 (2017).
- [19] Carroll, E. A. and Latulipe, C.: The creativity support index, *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, New York, NY, USA, Association for Computing Machinery, pp. 4009–4014 (2009).
- [20] Cherry, E. and Latulipe, C.: Quantifying the Creativity Support of Digital Tools through the Creativity Support Index, *ACM Trans. Comput.-Hum. Interact.*, Vol. 21, No. 4 (online), DOI: 10.1145/2617588 (2014).

^{*15} IK Multimedia - AmpliTube 4 <https://www.ikmultimedia.com/products/amplitude4/>

^{*16} Guitar Amp Simulator and Modeler Software — BIAS Amp 2, <https://www.positivegrid.com/bias-amp/>