アナウンス技術向上のための練習支援システムの試作

山田　真那子[1,a)]　塚田　浩二[1]

概要：
本研究では，音声情報処理技術を用いたアナウンスの練習支援システムを提案する．アナウンスとは一般的にテレビやラジオなどでアナウンサーが情報を分かりやすく伝えるための技術である．本研究では，高校や大学で行われる競技アナウンスの採点基準を参考にして，アナウンス技術の評価指標を考察する．さらに，アナウンス初心者／経験者の音声データを収集し，評価指標に基づく評価技術を実装して，アナウンス技術の定量的な分析を図る．さらに，分析結果をユーザに提示するフィードバックシステムを試作し，評価実験を通してその使いやすさ等を検証する．

Prototype of a Practice Support System for Improving Announcement Techniques

MINAKO YAMADA[1,a)]　KOJI TSUKADA[1]

1. はじめに

アナウンスは，聞いている人に情報を分かりやすく伝えるための技術であり，テレビやラジオのアナウンサーに必須とされるだけでなく，高校や大学でもアナウンス技術を競う大会がある [1][2]．アナウンスの練習では，熟練者のアナウンス音声を聞いて真似したり，熟練者からアドバイスを受けることが一般的である．しかし，熟練者のアドバイスは抽象的な表現を使った感覚的なものが多く，初心者が正確に理解することは難しい．例えば，初心者に対するアドバイスとして多いのが「文にうねりがあったから，しっかり下そう」というものがある．これは「文の音を一定に下げていかなければいけないのに，途中で音が上がってしまっていたので音を下げよう」という意味である．途中で音が上がるという意味の「うねり」や，文の音を下げながら読む「下し」などの専門的な表現が，初心者にとっては難しい．また，文章内のどの部分でどの程度音を下せていないのかも理解しづらい．そのため，初心者が感覚を掴めるようになるまで膨大な時間をかけて，熟練者と感覚をすり合わせるような繰り返しの練習が必要になる．さらに，一人で練習をすると無意識に悪い癖がついてしまうことがあり，こまめに熟練者に練習を聴いてもらう環境が必要になる．

本研究では，アナウンス技術を要素毎に分類・定量化し，学習者のアナウンスを分析／フィードバックすることで，一人でのアナウンス技術習得を支援するシステムの構築を目標とする．

2. 関連研究

本章では，本研究に関連する研究事例を，「歌唱や発話の支援」，「音声素材の音響的分析」，「ツールを用いた発話支援」の 3 つの観点から説明する．

2.1 歌唱や発話の支援

歌唱支援に関する事例として，中野ら [3] は，ユーザの歌唱力向上を支援するインタフェース Mirusinger を提案している．既存の楽曲のボーカルパートの音高（F0）を分析・可視化し，ユーザ自身の歌声と比較表示することができる．また，羽賀ら [4] は，歌唱の自主練習における学習効率の向上を目的とし，習熟度に関係する音響特徴量を可視化することで，学習者の歌唱データを評価した．プレゼ

1 公立はこだて未来大学
Future University Hakodate
a) g2124053@fun.ac.jp

© 2026 Information Processing Society of Japan

ンテーションに関する事例として，栗原ら [5] は，音声認識／画像認識を用いて，プレゼンテーション中の発表者の様子を分析し，リアルタイムにフィードバックするシステムを提案している．具体的な指標としては，話速度・声の抑揚・聴衆とのアイコンタクト等を用いている．また，趙ら [6] は KINECT センサを利用し，身体表現の「アイコンタクト」と「体の向き」，音声表現の「声の大きさ」を支援している．例えば，発表練習における良くない状態の時間が累積した場合，イヤフォンから音声で「前へ」，「大きい声」などと再生される．これらの研究と比較して，本研究では競技アナウンスの審査基準に基づいた分析やフィードバックを試みる点が特徴である．

さらに，声優の発話支援として，滑舌練習を行える「声優滑舌アプリ」[7] や，アフレコの練習を行うことが出来る「SAY-U」[8] などがある．一方，これらは音声分析等の知的機能は搭載されていない．また，声優では演技が重視されるのに対して，アナウンスは自然な発声で情報を正確に伝えることが求められるため，適した支援手法は変わってくると考える．

2.2 音声素材の音響的分析

王 [9] は，ナレーション（映像の説明音声），アナウンス（原稿の忠実な読み上げ），声優（声による演技）の 3 種類の音声データを対象として，音響的分析を行っている．例えば，声優はアナウンサーより母音の発音が長い．しかし，文末の句点による静止時間については三者とも長く，2 秒ほどとっていると述べている．アナウンスの初心者にとって間（無音時間）をしっかりとるのが難しいとされるが，アナウンス以外のナレーション／声優も，熟練者はしっかりと間を取っていることが分かる．

2.3 ツールを用いた発声支援

川村ら [10] は，加速度センサを用いて口唇振動を計測し，LED ライトによる視覚的フィードバックを提供するチューブ型のフィードバックシステムを提案した．ツール使用により，発声時持続時間の延長，声域の拡大また，北村ら [11] は，割り箸を使用した状態の顔面の動きと声の大きさをリアルタイムにフィードバックする発話訓練システムを開発．音圧や基本周波数の変化幅が上昇することが分かった．

本研究でも，今後の展望としてツールを用いた基礎練習支援を検討している．フィードバックシステムと組み合わせることでより効果的な練習支援を目指す．

3. 提案

本研究では，アナウンス技術を要素毎に分類・定量化し，学習者のアナウンスを分析／フィードバックすることで，技術習得を支援するシステムの提案を目指す．

3.1 評価指標

評価指標は，NHK 杯全国高校放送コンテスト [1] や NHK 全国大学放送コンテスト [2] におけるアナウンス部門の評価基準を参考に検討する．表 1 に，既存の評価基準と本研究での評価指標を整理したものを示す．次に，各評価指標について詳しく説明する．

3.1.1 発声

評価基準としての発声には，音量の一定性と語尾の発声の仕方の要素がある．音量の一定性は，声が大きすぎたり，小さすぎないように発声しているかという基準である．アナウンスでは常に一定の音量での発声が必要とされる．アナウンスは感情表現を激しくしてはいけないとされており，言葉を強調したい場合にも音量を大きくする表現方法は望ましくない．語尾の発声の仕方は，例えばよくある語尾の「～します.」という表現があった場合に，「しますぅ.」という風に最後の一音の母音が残ってしまう場合や，反対に「しますっ.」といった風に最後の一音を勢いよく発声してしまい，語尾に促音が入っているように聞こえてはいけないということである．

本研究では，より影響が大きい要素であると考え，前者の「音量の一定性」に着目する．初学者の特徴として，文頭の音を勢いよく発声してしまうため，文章の一音目が強くなってしまいがちである．また，アナウンス原稿の一文を読み切るのに，息が続かない・体力がないなどの要因から文末の音量が小さくなる傾向にある．これらの初学者特有の課題を解決するためにも，音量の一定性が評価指標として有効であると考えた．

3.1.2 発音

発音とは，いわゆる滑舌のことである．放送コンテストなどにおいて求められる滑舌は，日常会話のように「聞き取れれば問題ない」といった基準ではなく，音が濁ったり曖昧になったりしていないかを注意深く評価される．特に，サ行・タ行・ラ行は舌や口の動きが複雑であるため，滑舌の不明瞭さが表れやすい音である．

本研究では，この滑舌の明瞭さを音声認識による文字起こしされた原稿と，本来の原稿の差異を確認することで分析する．

3.1.3 アクセント

アクセントは個々の語について，社会的慣習として決まっている相対的な音の高低で表現される．例えば，「橋」・「箸」・「端」という全てが「はし」と読まれる単語でも，音の高低によって異なる意味を持つ．アクセントは通常，日本語発音アクセント新辞典 [13] などのアクセント辞典によって単語ごとに指定されており，音の下がる位置や鼻濁音，無声音の指示が示されている．また，アクセントには第一アクセント，第二アクセントなど一つの単語による複数のアクセント表記があるものもある．これらはどちらで読んでもいいとされる．アクセントを適切に読むことで正しい情報を相手に伝えられると同時に，間違えたアクセントによる違和感を無くし，聞き逃しを予防する．

本研究では，単語ごとの基本周波数（F0）を分析し，正しいアクセントで読まれているかを検証する．

3.1.4 イントネーション

イントネーションとは，音高の変化によって文意の理解に影

表 1 放送コンテストの審査基準と本研究の評価指標

審査基準	補足説明	本研究における評価指標
発声	音量の一定性，語尾の発声の仕方（伸びていないか，促音が入っていないか）	RMS の標準偏差等を用いた音量の一定性
発音	滑舌（一音一音の明瞭さ）	音声認識された原稿と，本来の原稿に差異が無いか
アクセント	単語の音の高低の正しさ	基本周波数（F0）に基づく，単語の音の高低の正しさ
イントネーション	一文の音が適度に下がっていくか	句読点で区切った場合に，基本周波数（F0）が適度に下がっていくか
テンポ	文章を読む速度	1 分間あたり 300 字程度で読めているか
ポーズ	句読点での間（無音時間の長さ）	句点での間は 2 秒，読点での間は 1 秒程度とれているか
言葉の立て方	重要な言葉を強調できているか（強調したい言葉の速度を遅くする，音を高くする，直前に間を入れる）	（未対応）

響を与える要素である．アナウンスにおいては，一般に文頭で音高を高く取り，文末に向かって音高を低下させる発話が基本とされている．

アナウンスでは，文意や係り受けの関係を明確にするため，文の途中で意図的に音高を上げる場合があり，これを「上げなおし」と呼ぶ．「上げなおし」を行うため，常に音高を単調に下降させていけばいいとは限らない．また，不適切な箇所で音高を上げてしまうことを，「うねり」と表現する．この「うねり」が発生してしまうと，文章の意味が正しく伝わらない可能性がある．文章の意味を正しく伝えられているか検証するため，文ごとの基本周波数（F0）の下がり方を分析する手法が有効であると考えた．

本研究では，句読点によって区切られた各句について，発話開始部と終了部の基本周波数（F0）を比較し，句末に向かって音高が低下しているかどうかを分析する．

3.1.5 テンポ

審査基準では文章を読む速度の具体的な指定はないが，聞き取りやすい速度で読むことが必要とされている．元 NHK アナウンサーであり，長崎大学でスピーチ研究を行っている矢野 [14] らは，文章を読んだときに聴き手が聞き取りやすい速度は，1 分間あたり 300 字程度としている．

本研究では，この値を基準の一つにするほか，経験者のアナウンス音声のテンポも参考にする．また，著者のアナウンス経験から，原稿全体を通した一定性だけでなく，句単位で重要度等を考慮した速度調整も必要と考えるため，こうした機能も検討する．

3.1.6 ポーズ

審査基準では具体的な間（無音時間）の秒数，箇所の指定はないが，聞き取りやすくなる間を取ることが必要とされている．一般的に間は句読点で取ることが多く，その長さは文章の重要度などを考えて一定範囲で調整する必要がある．

本研究では，2.2 節で紹介した王 [9] の研究や著者の経験を基に，目安として句点で 2 秒，読点で 1 秒程度の間をとることを当面の方針とする．

3.1.7 言葉の立て方

アナウンスでは，文章を分かりやすく伝えるために，情報の優先度をつけて読む必要がある．そのため，固有表現（固有名詞や，数詞など）など原稿の重要な部分を強調する．“遅く読む” というのは比較的使用しやすい技法で，シンプルに強調したい単語をゆっくり読むことである．“高く読む” は，強調したい単語の 1 音目で少しだけ音を上げて，また音を下げながら文章を読んでいくことである．“前に間を入れて読む” というのは，前述したポーズよりは短めの無音時間を入れることで言葉を強調する技法である．

言葉の立て方については，他の評価指標の複合的なものとなるため，本論文では対象とせず，今後の開発課題とする．

3.2 システム構成

システム構成図を図 1 に示す．システムは大きく分けて，「アナウンス技術の評価システム」，「GUI を備えたフィードバックシステム」の 2 つがある．評価システムは，各評価指標に基づき音声認識と音響分析を組み合わせた評価技術を開発する．初心者や大会成績上位者のデータを収集して検証することで，正しく評価できているかどうかを確認する．著者自身のアナウンス経験も生かして，適切な評価技術の開発を図る．フィードバックシステムは，適切な GUI を開発することで評価結果を初心者に分かりやすく伝える．

© 2026 Information Processing Society of Japan

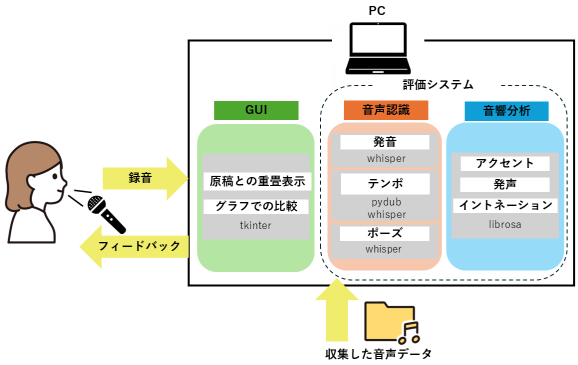

図 1　システム構成図

4.　実装

本章では，音声データの収集や前処理，評価システム，フィードバックシステムの実装の詳細について述べる．

4.1　データ収集と前処理

大学内の放送サークルのアナウンス初学者・経験者のアナウンスをマイクで録音し，アナウンス音声を収集する．現状では，アナウンス初学者 3 人，アナウンス経験者・熟練者 7 人のデータを収集している．

収集した音声データは，競技アナウンス特有の番号・氏名の読み上げや，録音環境の差による音量のばらつきが含まれる．そのため，以下の前処理を実施する．

- 音量正規化（ピークレベルを −0.1 dBFS に統一）
- 番号・氏名部分のトリミング
- 句読点単位のセグメント生成（音声区間と原稿位置の対応付け）

入力としては，アナウンス音声ファイルと対応する原稿ファイルを用意する．音量正規化および句読点単位のセグメント生成は自動処理により行う．一方，番号・氏名部分のトリミングについては，音声内容に依存するため，一部手動操作を含む．具体的には，音声から文字起こしした冒頭約 20 文字をシステム上に表示し，利用者がトリミングしたい箇所までの文字位置を指定することで，指定位置以前の音声区間を削除する．音声正規化は soundfile や librosa ライブラリで実装しており，出力結果は WAV（サンプリング周波数 48 kHz，ビット数 16 bit）で保存される．

また，熟練者のアナウンス音声については，過去の大会における上位入賞者の音声が CD や YouTube 上で公開されているため，これらの既存音声データも併せて利用した．これらの音声データは原稿が公開されていないため，Whisper ライブラリを用いて音声から原稿の文字起こしを行い，その後，収録データと同様の前処理を行って分析に利用している．

4.1.1　原稿セグメント

本研究では，Python を用いて原稿と音声の対応付けに基づくセグメント生成を行う．音声認識には Whisper を用い，音声から取得した単語単位のタイムスタンプ情報と，原稿を句読点単位で分割したテキスト情報を照合することで，各句に対応する音声区間を推定する．

セグメント生成については，原稿情報を基準として音声区間の分割および対応付けを行う．以降では，原稿を句読点（「。」「、」）で分割して得られる単位を「原稿セグメント」と定義する．原稿セグメントを音声区間と対応付けるために，音声認識結果に含まれるタイムスタンプを利用する．

具体的な手順を示す．まず，原稿を句読点（「。」「、」）で分割し，各句を正規化（空白除去，全角・半角統一）する．次に，Whisper の単語タイムスタンプを単語列として展開し，原稿セグメントと単語列の類似度を Python の difflib モジュールを用いて評価し，最も整合する範囲を探索する．さらに，探索により得られた開始・終了時刻を原稿セグメントに付与し，後述する評価指標（RMS，テンポなど）と共に JSON 形式で保存する．

Listing 1 に，JSON 形式の原稿セグメントの具体例を示す．ここでは，以下のような要素が含まれている：(1) 音声ファイルパス，(2) 各句の一意な識別子（例：s001），(3) 句のテキスト，(4) 音声中の開始時刻・終了時刻（秒），(5) 評価指標．

Listing 1　句読点単位セグメント情報の例

```
{
 "audio_path": "sample_audio.wav",
 "segments": [
  {
   "id": "s001",
   "text": 三年前に，"",
   "start": 0.5,
   "end": 1.2,
   "metrics": {
    "rms_ratio": 0.88,
    "tempo_ratio": −0.09,
    "mora_per_sec": 4.2
   }
  },
  {
   "id": "s002",
   "text": 岩手県陸前高田市で"",
   "start": 1.3,
   "end": 2.8,
   "metrics": {
    "rms_ratio": 1.05,
    "tempo_ratio": 0.15,
    "mora_per_sec": 5.8
   }
  }
 ]
}
```

© 2026 Information Processing Society of Japan

4.2 評価技術

評価指標に基づき，各指標を算出するための評価技術の開発を行う．本研究では，Python を用いた音声認識と音響分析により，各評価指標を算出する手法を提案する．

前述したように，入力として音声データおよび対応する原稿を与え，音声認識により句単位の開始・終了時刻を含む区間情報を取得する．取得した区間情報に基づき，各区間または全体に対して評価指標を算出する．具体的には，「発音」，「テンポ」，「ポーズ」には音声認識を用いて，「発声」，「アクセント」，「イントネーション」は音響分析を用いて分析する．

4.2.1 発声

発声は，RMS（Root Mean Square）を用いて算出する．RMS は瞬間的な最大値（ピーク）に比べて外れ値の影響を受けにくいという特徴を持つ．そのため，録音中の一時的な破裂音やノイズに左右されにくく，区間全体としての発声の大きさ・ばらつきを評価するのに適している．音声をフレームに分割し，フレームごとの RMS を求めることで，時間変化を含む音量の指標が得られる．本研究の実装では，RMS 計算の窓長（フレーム長）を 2048 サンプル，フレーム間隔を 512 サンプルとした．無音の影響を抑えるため，閾値以下の無音フレームを除去した後に RMS を算出し，平均・標準偏差・中央値などの統計量を出力する．また，音量の安定性を表す量として，文章全体の RMS の変動係数（標準偏差/平均）を算出する．

さらに，原稿セグメント毎の分析を行うために，RMS 比率（`rms_ratio`）という評価指標を算出する．これは，ある句における平均 RMS を音声全体の平均 RMS で正規化した相対的な指標である．

4.2.2 発音

Python のライブラリ Whisper を利用して，アナウンス音声の文字起こしを行う．学習者が読む原稿データと文字起こしデータの差異を検証し，発音の明瞭さを評価する．発音が明瞭である程，差異は少ないと考えられる．なお，書き起こしたものが同音異義語になってしまう場合があるが，アクセントが同じ単語の場合は文字起こし側の問題の為，差異としない．

4.2.3 アクセント・イントネーション

Python のライブラリ librosa を利用し，周波数分析を用いた音の高低について分析を行う．1 分半のアナウンス音声のピッチ (F0) を可視化した例を図 2 に示す．

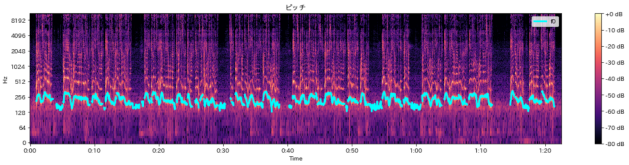

図 2 ピッチ抽出したアナウンス音声の例

現時点では，アクセントは単語ごとに NHK 日本語発音アクセント新辞典 [13] を用いて，手動で指定している．今後は，日本語音韻データベース*1 等を活用した実装を進めていきたい．

*1 https://www.cjk.org/language/ja/data/japanese/nlp/japanese-phonetic-database/

イントネーションは文章の音が下がっていくかを評価する．アナウンスでは，文頭の音が高く，文末にいくにつれ徐々に音を下げていくのが基本とされている．そのため，本研究では文章ごとに句読点で句切った時に，文末の音が文頭より下がっているかを分析する．

4.2.4 テンポ

テンポは，原稿に基づくモーラ数*2 と発話時間から算出する．原稿テキストは pykakasi を用いてひらがなへ変換し，ひらがな文字列からモーラ数をカウントする．発話時間は，音声から無音区間（ポーズ）を除去したうえで算出し，発声区間のみの総時間とする．さらに，セグメント生成時に得られるタイムスタンプを用いて，句ごとの発話時間を算出する．最後に，モーラ数を発話時間で除すことで， mora/sec を求める．この設計により，テンポの指標がポーズの長短に影響されにくくなり，発話そのものの速度を評価している．

4.2.5 ポーズ

句読点における無音時間の長さを分析する．原稿セグメントのタイムスタンプから，句で区切った際の無音時間を検出し，原稿データの句読点と対応した位置で適切に間を入れているかを評価する．

5. フィードバックシステム

フィードバックは手動調整を含まない実装が完了している評価指標（発声，テンポ，ポーズ）を対象として実装した．システムは，4.1.1 節で述べた原稿セグメントの生成と，生成したセグメントに基づく GUI 提示の二つの処理から構成される．前者により原稿上の位置と音声区間を対応付け，後者によりアナウンス音声の評価結果を原稿・波形・評価値とともに提示する（図 3）．

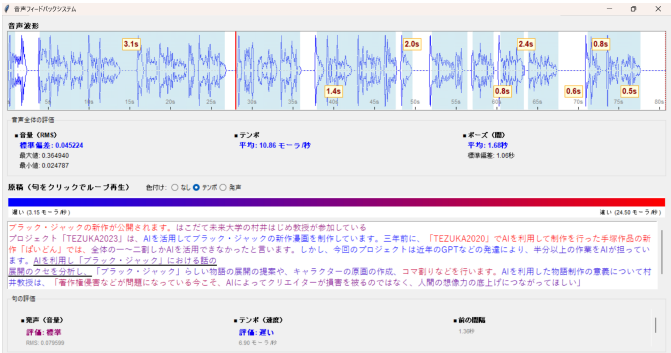

図 3 フィードバックシステム画面

5.1 フィードバック GUI

フィードバック GUI では，セグメント JSON と音声を読み込み，原稿と音声を同期させたループ再生と評価表示を行う．GUI は，「波形表示」，「原稿表示」，「評価表示」から構成される．

波形表示では録音した音声波形に重ねて，セグメントごとの

*2 モーラとは，日本語における発話の時間的単位である．仮名 1 文字（例：「か」「た」）を基本とし，「ん」「っ」「ー」なども 1 モーラとして数えられる．

© 2026 Information Processing Society of Japan

区間を水色でハイライトしている．原稿上の句をクリックすると該当区間のループ再生を開始し赤い縦線の再生ヘッドが再生中の時刻に対応する位置を動く（図 4）．

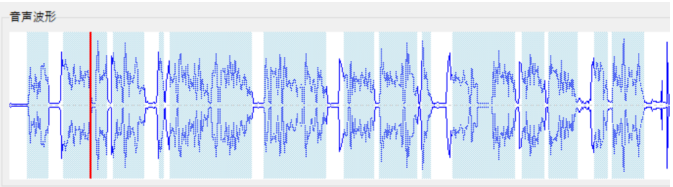

図 4　波形のハイライト表示と再生ヘッド

評価表示では，音量，テンポ，ポーズの各評価指標の分析結果が表示される（図 5）．

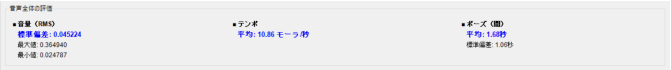

図 5　全体評価表示

原稿表示の部分は，各句をクリックすることで対応する音声区間をループ再生できる．また，ラジオボタンを選択することで，選択した評価指標に基づき，各句がグラデーション（赤から青）で色付けされる．この色付けは上部に配置したグラデーションバーと対応しており，音声ファイル内の評価結果の分布傾向を句単位で可視化できる（図 6）．

図 6　原稿表示

6.　ユーザ評価

6.1　目的

フィードバックシステムのユーザビリティを評価し，システムの改善点を調査する．

6.2　手法

被験者はいずれもアナウンス未経験者である大学生 11 名（男性 8 名，女性 3 名）とした．

実験手順は，まず実験概要を被験者に説明した後に，参考にしたい経験者の音声（3 名のアナウンス経験者が同じ原稿を読んだ内容）を被験者が任意に選択する．そして，被験者に 1 分半のアナウンス原稿を読み上げてもらい録音する（図 7）．この音声と選択した経験者の音声を分析した二つのフィードバック画面を並べて表示させ，比較してもらう．なお，操作の不明点などは随時実験者に確認できるものとした．最後に SUS や自由記述等のアンケートに回答してもらう．

ブラック・ジャックの新作が公開されます。はこだて未来大学の村井はじめ教授が参加しているプロジェクト「TEZUKA2023」は、AI を活用してブラック・ジャックの新作漫画を制作しています。三年前に、「TEZUKA2020」で AI を利用して制作を行った手塚作品の新作「ぱいどん」では、全体の一〜二割しか AI を活用できなかったと言います。しかし、今回のプロジェクトは近年の GPT などの発達により、半分以上の作業を AI が担っています。AI を利用し「ブラック・ジャック」における話の展開のクセを分析し、「ブラック・ジャック」らしい物語の展開の提案や、キャラクターの原画の作成、コマ割りなどを行います。AI を利用した物語制作の意義について村井教授は、「著作権侵害などが問題になっている今こそ、AI によってクリエイターが損害を被るのではなく、人間の想像力の底上げにつながってほしい」と話します。AI の力を借りて帰ってくる。「ブラック・ジャック」の新作は二〇二三年公開予定です。

図 7　アナウンス原稿の例

6.3　実験結果

SUS の結果は平均値が 67.5 と標準的水準であり，標準偏差が 18.0 とばらつきが大きかった．

次に，SUS の奇数設問と偶数設問の結果を図 8 と 9 に示す．グラフの上部に行くほど高評価を示す．奇数設問では Q5（機能の一貫性）と Q7（学習の容易さ）の評価がそれぞれ平均 4.1，4 と特に高い．自由記述のアンケートでも，機能の使い方への疑問は少なかったためツールの機能への理解は一定以上できていたと考えられる．また，偶数設問では Q4（専門的なサポートの必要性）と Q10（必要な事前知識の量）は平均 3.27，3.36 とやや低い．自由記述アンケートでも，アナウンスの事前知識が無いため，どの機能を使えば経験者と自分との差が確認できるのか分かりづらいという意見が出ていた．

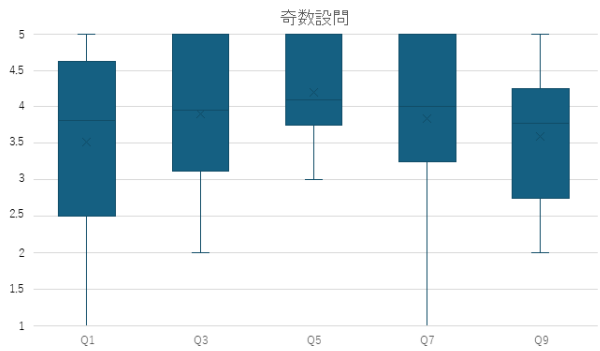

図 8　SUS 奇数設問の評価結果．1: negative 〜 5: positive.

© 2026 Information Processing Society of Japan

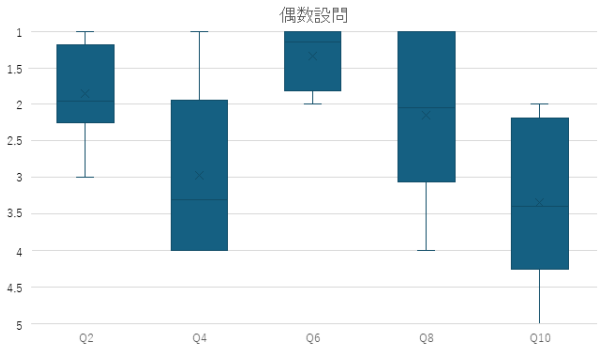

図 9 SUS 偶数設問の評価結果．1: positive ～5: negative.

7. 議論

7.1 フィードバックシステムの改善

ユーザ評価の結果を受け，システムの改善を行った（図 10）．一つの画面で経験者と自分の音声の比較を行えるよう変更した．音声波形は縦に並べ，ポーズの時間は直接比較できる．原稿は一つのみ表示し，句を左クリックで経験者，右クリックで初心者，ダブルクリックで両者の音声を同時再生できる．また，句ごとの評価結果について，経験者と初心者の結果を右上に同時に表示した．さらに，テンポを例とすると，経験者の値と比較して「非常に速い」「速い」「適正」「遅い」「非常に遅い」のいずれかをラベルとして表示し，その差分も併せて下部に表示した．

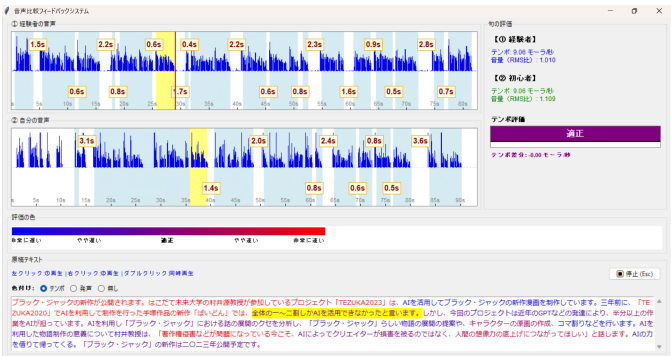

図 10 改善後の GUI 画面

7.2 基礎練習の支援

フィードバックシステムと組み合わせて，直接的なアプローチの技術向上支援を今後行っていきたい．現在，発声・発音を改善する割り箸を使った発声練習に注目している．

割り箸を使った発声練習は，アナウンスだけでなく歌唱力向上を目的としても活用されている．割り箸発声練習を行うことで，口の開き方が正しい形になることや，顎の動きが制限されるため顎以外の筋肉が鍛えられる．しかし，指導者により箸の本数やくわえ方にばらつきがあるため，アナウンスに合った練習方法の模索を行う．

本研究ではツールにセンサを取り入れた上で，フィードバックシステムと連携し，練習前後の変化を視覚的に提示することを検討している．

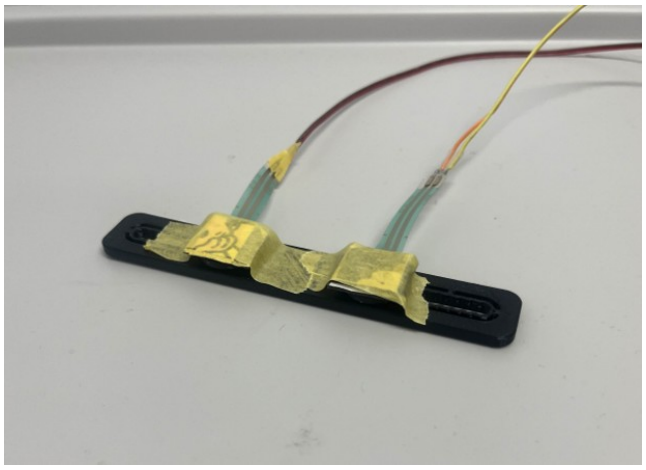

図 11 割り箸型発声練習支援ツール

現在試作中のツールは，圧力センサを左右二つに設置して，奥歯で噛みながら発声練習を行うことで，常時噛む力を計測する（図 11）．マイコンは Arduino UNO を使用しセンサ値を取得し，PC に保存する．筐体部分は 3D プリンタで製作し，ラップなどでくるんで利用する．

8. まとめと展望

本研究では，高校や大学で行われる競技アナウンスの採点基準を参考にして，アナウンス技術の評価指標を考察した．次に，アナウンス初心者／経験者の音声データを収集し，評価指標に基づいて，アナウンス技術を分析するシステムと，発声とテンポを対象としたフィードバックシステムの実装を進めた．さらに評価実験では，フィードバックシステムについてのユーザ評価を行い，その結果に基づいてシステムの改善を行った．

今後，フィードバックシステムに他の評価技術の組み込むとともに，割り箸型発声練習支援ツールとの統合を進める．さらに，これらの改良を施したフィードバックシステムを対象として，アナウンスの練習支援効果の評価実験の実施を目指す．

参考文献

[1] 第 70 回 NHK 杯全国高校放送コンテスト「要項号」，https://www.nhkfdn.or.jp/kyoiku/ncon/ncon_h/pdf/70/70_youkou.pdf，(参照 2023-06-25)．

[2] 第 40 回 NHK 全国大学放送コンテスト要項，https://drive.google.com/file/d/18lIMIX3z1oDr_XM5PwYRZQct0MLZxN8l/view?usp=drive_link，（参照 2023-06-25）．

[3] 中野 倫靖，後藤 真孝，歌声インタフェース：歌声を対象とした信号処理とそれに基づくインタフェース構築，日本音響学会 2013 年秋季研究発表会 講演論文集，3-7-3，pp. 1465-1468(2013).

[4] 羽賀 翼，香山 瑞恵，池田 京子，橋本 昌巳，伊東 一典，習熟度に関係する音響特徴量に基づく 歌唱学習支援システムの評価，人工知能学会第二種研究会資料，身体知研究会，SIG-SKL-20-01(2015).

[5] 栗原 一貴，後藤 真孝，緒方 淳，松坂 要佐，五十嵐 健夫，プレゼン先生：音声情報処理と画像情報処理を用いたプレゼンテーションのトレーニングシステム，WISS2006

© 2026 Information Processing Society of Japan

論文集，NO．43，pp59-64(2006).
[6] 趙 新博，由井薗 隆也，宗森 純，初心者向けプレゼンテーション練習支援システム PRESENCE の開発と評価，電気学会論文誌 C，Vol.138，NO.10，pp1269-1277(2018).
[7] IAM，K.K.，声優滑舌アプリ，https://apps.apple.com/jp/app/%E5%A3%B0%E5%84%AA%E6%BB%91%E8%88%8C%E3%82%A2%E3%83%97%E3%83%AA/id642099721，（参照 2023-10-30）.
[8] Space Factory Inc., SAY-U, https://apps.apple.com/jp/app/say-u/id1533211139,（参照 2023-10-30）.
[9] 王 伸子，日本語教育の教材としての音声素材の音響的分析 ―ナレーション、アナウンス、声優ボイスオーバーの分析― ，専修大学外国語教育論集，50 57-73,(2022).
[10] 川村 直子，チューブ発声中の口唇部の振動を可視化する携帯型フィードバックシステムの開発 ，音声言語医学，64 10-17，(2023).
[11] 北村 達也，顔面の動きをフィードバックとして用いる発話訓練のシステムの検討，日本音響学会講演論文集，1357-1358，(2019).
[12] 柴田 忠彦，コールセンターオペレーターの音声品質評価の自動化の研究，滋賀大学修士論文，2022.
[13] ＮＨＫ放送文化研究所，日本語発音アクセント新辞典，NHK 出版，2016.
[14] 矢野 香，【第 2 回】「ＮＨＫ式 7 つのルール」をマネれば、あなたの話し方が一変する！,DIAMOND online, https://diamond.jp/articles/-/56483?page=3, (参照 2024-01-23).

© 2026 Information Processing Society of Japan