口形に着目した発話訓練が音声に及ぼす影響*

☆設楽郁巳, 北村達也 (甲南大), 牧野桃子, 山根典子 (広島大)

1　はじめに

発話に関して医学的な所見が見られない大学生・大学院生の約 3 割が，日常的な発話に苦手意識を感じており，そのほとんどが発話の改善を希望していた [1]. 発話に対する苦手意識を抱えたままでは，面接や発表等の場面において実力を発揮できない可能性がある. 鴨居・沼田 [2] は発音，発声，滑舌などを学ぶことによって，人前で話すことに対する不安や心配，自信がない等のネガティブな意識が軽減されると報告している．これらの背景から，我々は簡便かつ効果的な発話訓練法について検討している.

明瞭性が高く聴き取りやすい発話を行うために，様々な発話訓練法が検討されている．本研究では，アナウンスの発話訓練にて口角を上げて発話するよう指導される点に着目した．発話時の口角の位置を上昇させることで発話が改善されると仮定し，発話時の口形に着目した発話訓練が音声に及ぼす影響について検討した.

2　発話時の口角位置調査

発話訓練経験の有無により発話時の口角位置に違いはあるのか確認した.

2.1　実験方法

実験参加者は大学の放送サークルに所属し発話訓練経験のある大学生 6 名 (発話訓練経験あり群)，発話訓練経験のない大学生 10 名 (発話訓練経験なし群)，計 16 名である．実験参加者がイソップ童話「北風と太陽」の文章を読み上げている最中の顔の動画をビデオカメラ (Panasonic HC-V360MS) により撮影した.

2.2　口角位置判定

収録した動画のうち発話区間のみを取り出し調査を行った．口角の位置の判定には，顔ランドマーク座標を推定する機能がある機械学習ライブラリ MediaPipe の Face Landmark Detection を使用した．左右の口角と，上唇中央の最下部，下唇中央の最上部の 4 点の座標を自動的に抽出し口角位置の判定に利用した. 上唇と下唇を結んだ線の中点の座標を計算し，その中点が左右の口角を結ぶ線の上下どちらに存在しているか計測した．図**??**のように，左右の口角を結ぶ線が，上唇と下唇を結んだ線の中点より上に位置していれば口角は上位であると判定した．この判定法により，2 群の話者が発話時間全体の内，どれだけの間口角が上位だったか調査した．発話時間全体に対する口角が上位であった時間率のことを，本研究では口角上位時間率と呼ぶ.

2.3　結果

調査結果を図 1 に示す．発話訓練経験あり群では，平均して発話時間全体の 83.6%の間口角が上位と判定された．それに対し，発話訓練経験なし群では平均して 12.8%の間しか口角が上位と判定されなかった.

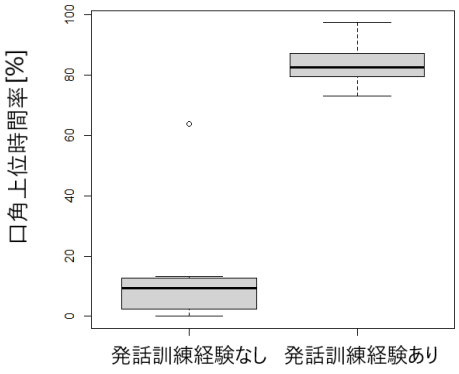

Fig. 1: 両群の口角上位時間率

3　口形に着目した発話訓練法の検討

2 節の調査から，発話時の口角の位置を上昇させることが発話の改善に繋がると仮定し，口形を用いた発話訓練支援システムを開発した．システムのスクリーンショットを図 2 に示す．システム利用者は PC の Web カメラに自分の顔を映し，口角の位置を確認しながら発話を行う．口角の上位判定には 2.2 節の手法を用いた．画面中央には Web カメラで撮影した利用者の顔が表示され，それに重なる形で左右の口角とそれを結ぶ線，上唇と下唇の中点の位置が表示される.

画面中央に四角形の枠が表示され，利用者はこの枠内に自分の顔が収まるよう意識する．これは利用者と Web カメラとの距離を最適に保つねらいがある. 利用者の口角の位置が下の場合，画面には「口角が下です」と表示し，枠の色を青にする．口角の位置が上の場合，画面には「口角が上です」と表示し，枠の色を赤にする．利用者は「口角が上です」の表示をで

*Effects of voice exercise focusing on mouth shape on speech. by SHITARA, Ikumi, KITAMURA, Tatsuya (Konan Univ.), MAKINO, Toko, YAMANE, Noriko (Hiroshima Univ.)

きるだけ長い時間維持するよう意識し発話を行う．画面上部には発話練習の残り時間と練習文が表示され，練習文はあらかじめ設定した時間ごとに切り替わる．

本システムは JavaScript および HTML により記述され，仮想サーバ上に配置した．利用者はインターネットに接続し Web ブラウザが使える環境さえあれば，本システムを利用できる．

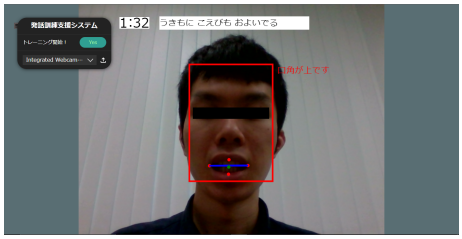

Fig. 2: 提案システムのスクリーンショット

4　実験

4.1　実験方法

実験参加者は 18 歳から 24 歳の大学生・大学院生 30 人 (男性 22 名，女性 8 名) で，3 節のシステムを用いる実験群 (15 名) と用いない統制群 (15 名) に分けた．発話資料は北原白秋「あめんぼ」の全文と，イソップ童話「北風と太陽」の文章である．

実験に先立ち，実験参加者に実験説明を口頭と文書で行い署名による同意を得た．実験参加者は椅子に座り，ディスプレイ (BenQ GW2270-T) に表示された発話資料を読み上げた．話者の顔の正面から約 0.5 m の位置にウェブカメラ (Logicool BRIO) を配置し，動画を撮影した．マイク (Sony ECM-77B)，オーディオインタフェース (Roland Rubix 24) を使用し音声収録も行った．

次に，実験参加者に自宅にて自己練習を行わせた．実験群には本研究で開発したシステムを使用させ，統制群には PC の画面中央に表示した文が一定間隔ごとに切り替わる別のシステムを 1 日 1 回，2 週間の間使用させた．1 回に要する時間は約 15 分程度である．実験参加者には PC スタンドを貸与し，練習時顔が下を向かないようにさせた．

自己練習開始から 1 週間及び 2 週間経過時に，実験参加者の顔画像の撮影と音声の収録を行った．その際，実験群には発話資料の読み上げ方法を「発話訓練支援システムで練習したように，口角を上げるよう意識して」と指示した．

収録した音声を刺激音とする聴取実験を大学の放送サークルに所属し発話訓練経験のある 4 名 (男性 1 名，女性 3 名) を対象に実施した．発話訓練前後の音声を連続して提示し，どちらが良い発話だと思ったか一対比較法により判断させた．

4.2　結果と考察

実験群，統制群の「北風と太陽」発話時の，口角上位時間率の遷移をそれぞれ図 3，図 4 に 表す．横軸は実験開始から経過した日数を表している．実験群では 8 日目以降の収録で口角上位時間率が向上する傾向が見られたのに対し，統制群では初回の収録時と 2 回目以降もほぼ変わらない傾向が見られた．

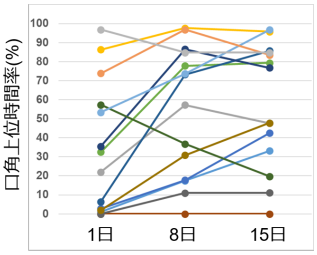

Fig. 3: 実験群の口角上位時間率

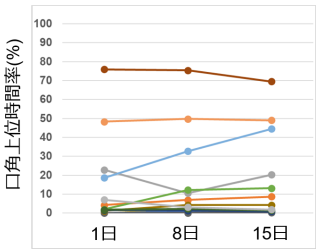

Fig. 4: 統制群の口角上位時間

聴取実験の結果を図 5，図 6 に示す．群ごとに分析した結果，実験群の訓練後の音声が統制群より高く評価された．

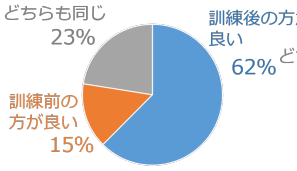

Fig. 5: 実験群

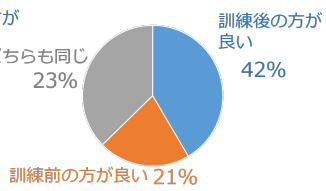

Fig. 6: 統制群

5　おわりに

本研究では，発話中の口角位置に着目し，その位置をフィードバックする発話訓練支援システムの開発と評価を行った．システムには口角を上げさせる効果が確認され，訓練後の音声の印象も向上する結果が得られた．現在，発話中の口角の位置だけでなく開口面積もフィードバックするシステムを開発中である．

謝辞　データ整理にご協力いただいた甲南大学知能情報学部 渡邉ひかりさんに感謝します．本研究は，中山隼雄科学技術文化財団，JSPS 科研費基盤研究 (JP20H00291) の助成を受けた．

参考文献

[1] 北村ら, 音響誌, 75(3), 118–124 (2019).
[2] 鴨居, 沼田, 日本音声学会全国大会予稿集,127–132 (2022).