



# リアルタイムな擬似安静声フィードバックによる 緊張緩和手法の提案

成瀬加菜<sup>1)</sup>, 吉田成朗<sup>1)</sup>, 世田圭佑<sup>1)</sup>, 鳴海拓志<sup>1)</sup>, 谷川智洋<sup>1)</sup>, 廣瀬通孝<sup>1)</sup>

1) 東京大学大学院情報理工学系研究科 (〒 113-0033 東京都文京区本郷 7-3-1)

**概要:** 本研究では, 口頭発表における緊張緩和を目的とした変換聴覚フィードバックシステムを提案する. このシステムでは緊張状態の声の特性を打ち消すように入力音声を漸次的に変換し, 擬似的な安静状態での声をフィードバックすることで, 発話の明瞭化を使用者に知覚させ, 緊張緩和を狙った. スピーチ実験による効果検証の結果, 音声パラメータの違いが緊張感に影響した可能性や, システムが使用者の自信向上に寄与した可能性が示唆された.

**キーワード:** 聴覚, 緊張感, スピーチ, 変換聴覚フィードバック

## 1. 研究の背景と目的

充実した人生を送る上で, 他者とのコミュニケーションは必要不可欠である. しかし, 他者を前にして話す状況では緊張感が喚起されやすい. 緊張感が生じる背景には失敗への恐怖心や聴衆の反応などが挙げられ [1], これらを過剰に意識することはスピーチパフォーマンスの低下や印象の悪化を招く場合もある [2]. 本研究では, 他者への意思伝達の円滑性向上を目的とし, 緊張感の原因となる感情情報の操作を行うことで, 発話活動における緊張感を緩和するシステムの構築を目指す.

## 2. 関連研究

### 2.1 変換聴覚フィードバックの概要と効果

本研究では, 発話活動と親和性の高い緊張緩和システムを構築するために, 感情に作用する工学的手法の一つである変換聴覚フィードバック (Altered Auditory Feedback, AAF) を利用する. AAF とは, マイクに入力された音声に変換処理を行い, ヘッドホンを通じて発話者自身にフィードバックする手法である.

AAF には使用者に特定の感情を生起させる効果が知られている. Aucouturier et al. は, 入力された発話音声に変換処理を行うことで, 喜び・悲しみ・恐れ各感情を表現した変換音声を生成するプラットフォームを開発した [3]. 検証実験では, 与えられた文章を参加者が朗読する間, 発話音声を上記の感情を表現した音声へと徐々に変化させてフィードバックした. その結果, 参加者にはフィードバック音声に表現された感情が実際に喚起されることが確認された. このことから, 人間は自身の声と感情とが合致しているかを意識するのではなく, 声に表れた感情を自身の感情と推測している可能性が示唆された.

本研究では AAF の感情作用を利用し, 緊張感による使用者の声の変化を音声変換で抑制することにより, 緊張感の緩和を狙う.

### 2.2 声と緊張感の関係

AAF が感情に作用する原理は, 人間の声と感情には関係性があり, 発話者の感情が声の聴覚情報から推測可能であることによって説明される. 緊張感も, 喉頭や声帯の筋肉を緊縮させることで発話者の声を変化させることが知られている.

Russell の円環モデルでは, 緊張状態は覚醒度の高い不快感情として, 反対に安静状態は覚醒度の低い快感情として位置付けられている [4]. 特に覚醒度と声の音響パラメータには強い相関があることが知られている. 覚醒度と正の相関を持つパラメータには, 基本周波数, 高周波数成分のエネルギー比率, 声の震えなどが挙げられ, 覚醒度の増大に応じてピッチの高い鋭利な声になる傾向がある [5]. 本研究では, 緊張状態の声に現れる上記の特性を音声変換で打ち消す AAF システムを提案する.

## 3. 緊張緩和を目的とした AAF システムの構築

発話者の緊張感を緩和するため, 緊張状態での声の特性を打ち消すように入力音声を変換し, 生成された擬似的な安静状態での声をフィードバックする AAF システムを構築した.

変換音声を発話と同時に出力する必要があるため, リアルタイムな変換が容易である基本周波数と周波数スペクトルのエネルギー分布の 2 つのパラメータに対し変換処理を行った. 処理の流れを図 1 に示す. まず, マイクへの入力音声にピッチシフトを行い, 基本周波数を低下させた. 次に, ゲインを負の値にしたハイシェルフフィルタを用いて高周波数成分の信号を減衰させ, 変換音声をヘッドホンから出力した.

著者らは本稿で後述する検証実験を行う前に, 上記手法

Kana NARUSE, Shigeo YOSHIDA, Keisuke SETA,  
Takuji NARUMI, Tomohiro TANIKAWA and Mi-  
chitaka HIROSE

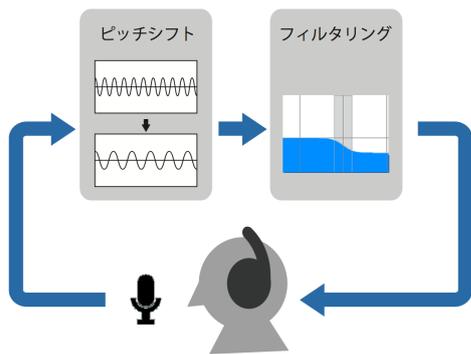


図 1: 音声入力から出力までの流れ

の効果を検証するため、AAF を用いたスピーチ実験を行った [6]。その結果、変換パラメータを変えることが使用者の緊張感に影響したことは確認されず、パラメータに関係なく変換音声への違和感が緊張感を増大させる可能性や、声の明瞭化を知覚することが緊張緩和に寄与する可能性が示唆された。この結果に基づき、AAF 使用開始直後には無変換の発話音声をフィードバックし、漸次的に変換を行う機能をシステムに追加した。この機能によって、変換音声への違和感が軽減され、パラメータの違いが緊張感に影響しやすくなると考えた。また、漸次的な音声変換により声が明瞭化する過程を使用者に知覚させることで、緊張緩和効果を得ることができると考えた。

#### 4. 緊張緩和効果の検証実験

擬似的な安静状態での声の AAF が緊張感に及ぼす影響を調査するため、検証実験を行った。実験では、希望職に就くための面接という場面設定のもとで、参加者は AAF を使用しながら 5 分間のスピーチを行った。

##### 4.1 実験条件と変換パラメータ

参加者は 19 歳から 32 歳までの 12 名（うち女性 7 名）で、平均年齢は 23.08 歳だった。面接官として参加者のスピーチを観察する実験協力者は 22 歳から 24 歳までの男性 3 名で、平均年齢は 22.67 歳だった。

実験は以下の 3 条件からなる参加者間計画で行った。

- 安静声条件：スピーチ開始時は無変換の発話音声をフィードバックし、スピーチ開始 60 秒後から 120 秒後にかけて徐々に緊張時の声の特性を抑えた音声へと変換した。以下はスピーチ開始 120 秒後以降の変換パラメータである。
  - － ピッチシフト：50 セント下降
  - － フィルタ：カットオフ周波数 1 kHz, オクターブ毎のゲイン-9.5 dB のハイシェルフフィルタ
- 歓喜声条件：Aucouturier et al. [3] が AAF による喜び感情の喚起に用いた音声変換と同様のものを使用した。スピーチ開始時は無変換の発話音声をフィード

バックし、スピーチ開始 60 秒後から 120 秒後にかけて徐々に喜びを表す音声へと変換した。以下はスピーチ開始 120 秒後以降の変換パラメータである。

- － ピッチシフト：50 セント上昇
- － フィルタ：カットオフ周波数 8 kHz, オクターブ毎のゲイン-9.5 dB のハイシェルフフィルタ
- － 抑揚：発話開始の各タイミングでピッチシフトを-170 セントから 170 セントまで急速に変化させ、その後 500ms の間に 30 セントまで徐々に下げる

- 無変換条件：スピーチ中、一貫して無変換の発話音声をフィードバックした。

上記全ての条件において、発話から音声フィードバックまでの遅延時間は 50ms 以下となるようにした。

参加者は実験開始の前日までに後述の特性不安検査 [7] に回答した。回答に基づき、性格特性として不安になりやすい程度を表す点数を各参加者について算出した。その後、点数の平均値が条件間で可能な限り均等になるよう、各参加者に 1 つずつ条件を割り当てた。

##### 4.2 実験手順

実験は大学内の一室で行った。実験部屋には机を 1 つと椅子を 3 つ設置し、机の上には PC を 2 台配置した。2 台の PC のうち 1 台は音声変換とアンケートの回答に使用し、マイク付きヘッドホンと接続した。もう 1 台の PC は皮膚電位計測器と接続し、緊張感の定量評価に使用した。

実験開始にあたり、参加者は実験に関する説明文書を読み、同意書への記入を行った。本実験の真の目的を知らせることが結果に影響することを防ぐため、参加者には実験目的を「スピーチ時の人間の振る舞いに関する調査」であると伝えた。

最初に、AAF に使用するヘッドホンの音量調整を行った。参加者は、ヘッドホンから自身の声が無変換でフィードバックされる状態で、事前に用意された文章を朗読しながら音量を調整した。音量は、実際の発話音声よりもヘッドホンからの出力音声の方がはっきりと聞こえる程度まで上げるよう、参加者に指示した。

次に、皮膚電位計測のための電極を装着した後、参加者は 3 分間目を閉じて安静状態をとった。その後、参加者は後述の状態不安検査 [7] に回答し、安静状態における緊張感の自己評価を行った。

安静状態終了後、スピーチの内容を考えるための準備時間を設けた。スピーチのテーマが提示された後、参加者は 5 分間の準備を行った。スピーチのテーマは全参加者に対し共通で、希望の職に就くための面接を想定した自己 PR とした。準備時間中、参加者はメモ用紙として A4 の紙 1 枚を自由に使用することができたが、スピーチ中にメモ用紙を参照することは許されないものとし、準備時間終了時に参加者から回収した。

スピーチ開始前に、面接官役を務める実験協力者 1 名が入室し、机を挟んで参加者と向き合うように着席した。参加者はその場で起立し、協力者と対面して 5 分間のスピーチを行った。スピーチ開始と同時に AAF を開始し、実験者は参加者の目につかない場所へ移動した。協力者は入室から退室まで、相槌、頷き、表情の変化を含む一切の反応をしないものとした。

スピーチ終了後、実験者は元の位置に戻り、協力者は退室した。参加者は状態不安検査に再び回答し、スピーチ中の緊張感の自己評価を行った。最後に、参加者は実験全体に関する質問やヘッドホンからの出力音声に関する質問を含むアンケートに回答した。

#### 4.3 仮説

先行研究では、特定の感情が喚起されていない状態で AAF を使用しながら文章を朗読することにより、変換音声に表現された感情が使用者に喚起されることが示された [3]。したがって、安静声条件の参加者は、自身が安静状態にあると知覚し、緊張感が軽減されると考えた。

歓喜声条件においても、無変換条件と比較して喚起される緊張感は軽減されると考えた。喜びは緊張と同じく覚醒状態であるが、不快感情である緊張に対し喜びは快感情であるため、緊張状態を快感情と再評価できる [8] 可能性があると考えた。

また、フィードバック音声を変換する 2 条件では、自身のパフォーマンスに対してポジティブな印象を持つことができ、無変換条件と比較してスピーチ後の自信度も向上すると考えた。

#### 4.4 評価方法

スピーチにより喚起される緊張感は、状態-特性不安検査 (State-Trait Anxiety Inventory, STAI) [7] と皮膚電位反応 (Skin potential reflex, SPR) によって評価した。

STAI は質問紙によって測定される不安尺度であり、性格特性として不安になりやすい程度を表す特性不安と、測定時の心理状態としての不安度を表す状態不安を独立して測定することが可能である。本実験では、安静状態終了直後とスピーチ終了直後に状態不安検査を 1 回ずつ行い、それらの点数の差分を求めることでスピーチにより喚起された緊張感を評価した。また、特性不安検査の点数をもとに参加者への実験条件割り当てを行った。

SPR は精神性発汗に由来する短期的な皮膚電気活動の一つであり、ストレスや興奮の指標として利用される [9]。本実験では、安静状態、スピーチ準備時間、スピーチ本番の各タイミングについて SPR の総量 (基線からの振幅の積算量) を求め、1 分間あたりの平均をそれぞれ算出した。

また、参加者は実験終了時のアンケートにおいて、スピーチ直後の自信の度合いを 7 段階で評価した。

#### 4.5 実験結果

状態不安検査について、スピーチ課題直後での点数から安静状態終了直後での点数を引いた差分を図 2 に示す。安静声条件 ( $M=14.0$ ,  $SD=7.68$ ), 歓喜声条件 ( $M=16.5$ ,

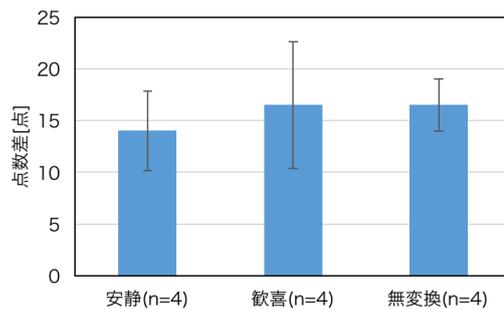


図 2: 状態不安検査の点数差 (エラーバーは標準誤差)

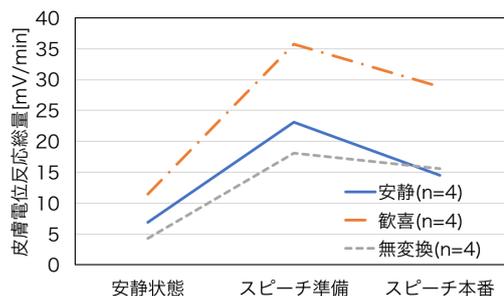


図 3: 皮膚電位反応量の推移

$SD=12.26$ ), 無変換条件 ( $M=16.5$ ,  $SD=5.02$ ) の 3 条件と比較し、クラスカル・ウォリス検定を行った結果、有意差は得られなかった ( $p = 0.8741$ , 効果量  $\eta^2 = 0.02448$ )。

SPR の推移を図 3 に示す。準備時間とスピーチ本番との反応量を比較すると、3 条件共に本番での反応量減少が見られた。

スピーチ中での SPR の推移を図 4 に示す。安静声条件と無変換条件では時間経過と共に反応量が減少したが、歓喜声条件では常に高い反応量を記録する傾向が見られた。

実験参加者のスピーチ課題終了時の自信度を図 5 に示す。安静声条件 ( $M=3.75$ ,  $SD=0.74$ ), 歓喜声条件 ( $M=2.25$ ,  $SD=0.22$ ), 無変換条件 ( $M=2.25$ ,  $SD=0.22$ ) の 3 条件と比較し、クラスカル・ウォリス検定を行った結果、有意差は得られなかった ( $p = 0.1612$ , 効果量  $\eta^2 = 0.3319$ )。しかし、安静声条件では他の 2 条件と比較して高い自信を得られる傾向が見られた。

#### 4.6 考察

スピーチ中における SPR 推移 (図 4) を見ると、安静声条件と無変換条件では時間経過と共に SPR が減少傾向を見せたのに対し、歓喜声条件では増加傾向が見られた。この結果は歓喜声条件で緊張感が軽減されるという仮説と反するが、歓喜声条件ではピッチ上昇や高周波数成分の信号増幅を行ったため、緊張状態での声に現れる特性が強調されたことが原因であると考えられる。実際に、声の音響パラメータは快感情よりも覚醒度との相関が強いことが知られている。そのため、緊張感の喚起されやすいスピーチ場面では、変換音声から知覚された覚醒度が緊張感に原因帰属された

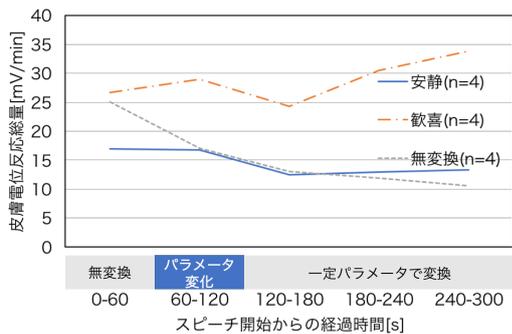


図 4: スピーチ中での皮膚電位反応量の変化

質問：次にスピーチやプレゼンテーションなどをする機会があった場合、自分はきちんと話せると思いますか？

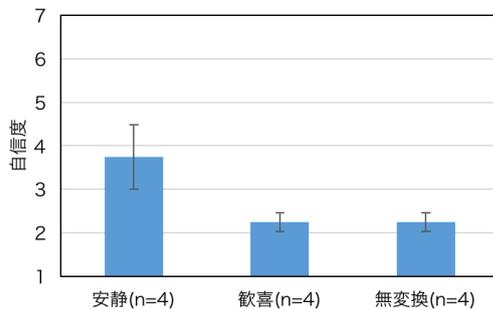


図 5: スピーチ課題終了時の自信度 (エラーバーは標準誤差)

ことで、参加者の緊張感が増大された可能性がある。

状態不安検査や実験全体での SPR 推移は 3 条件間に大きな差は見られず、スピーチ中の SPR 推移の結果からも、無変換条件と比較して安静声条件が緊張緩和効果を有することは確認できなかった。そのため、本実験における擬似安静声を生成する手法が有効であったと結論づけるためには更なる検証を必要とする。しかし、安静声条件と歓喜声条件との比較により、変換パラメータの違いがスピーチ中の緊張感に影響を与える可能性は示唆された。

また、安静声条件の参加者が他の 2 条件と比較してスピーチ直後の自信度を高く評価する傾向が見られたことから、提案手法が参加者の自信喚起に寄与した可能性がある。実験後の自由記述式のアンケートでは、安静声条件において、明瞭に発話している感覚を知覚できたことが自信に繋がったという回答を得た。一方、歓喜声条件では不安に感じられる声を聞くことで一層不安になったという回答があり、フィードバック音声の聞こえ方が AAF 使用者の自信に影響する可能性が示唆された。

## 5. 結論と今後の展望

本研究では、口頭発表場面での緊張緩和を目的とした AAF システムを構築し、スピーチ実験によって効果検証を行った。緊張感の生起により、声の基本周波数や高周波数成分のエネルギー分布が増大するという知見から、これらの変化

を抑制する音声変換を入力音声に対して行うことで、擬似的な安静状態での声を生成しフィードバックするという手法を提案した。実験の結果、提案手法が緊張緩和効果を持つことは示されなかったが、変換パラメータの違いが AAF 使用者の緊張感に影響する可能性が示唆された。また、AAF により自身の声が明瞭化した感覚を使用者に提示することで、スピーチに対する自信が生起される可能性が示唆された。

今回行った実験では安静声条件と無変換条件の緊張緩和効果に大きな差が見られなかったため、今後は変換手法の再検討を行う必要がある。変換音声に対する印象の調査などを行うことで、安静状態や自信を持った状態での声のパラメータを推定し、より効果的な音声変換の実現を目指す。また、実際にシステムが使用者に利益をもたらすものであるか検証するため、スピーチパフォーマンスの変化を評価指標に含める必要があると考える。

本研究で提案したシステムが緊張緩和を実現できれば、準備を必要としない即時的な感情制御が可能となるため、事前対策の難しい相互コミュニケーションにも応用が期待できる。今後はシステムの改善や、より精密な検証を行い、緊張緩和とスピーチパフォーマンスの向上を可能とするシステムの構築を目指す。

## 参考文献

- [1] 有光興記:「あがり」のしろうと理論:「あがり」喚起状況と原因帰属の関係, 社会心理学研究, 17.1, pp. 1-11 (2001).
- [2] 岩田彩香 他: 社交不安傾向によるスピーチ場面でのパフォーマンス低下に関する検討, 早稲田大学臨床心理学研究, Vol. 15, No. 1, pp. 53-63 (2015).
- [3] Aucouturier, J.-J. et al.: Covert digital manipulation of vocal emotion alter speakers' emotional states in a congruent direction, Proceedings of the National Academy of Sciences (2016).
- [4] Russell, J. A.: A circumplex model of affect., Journal of personality and social psychology, Vol. 39, No. 6, pp.11-61 (1980).
- [5] Juslin, P. N. and Scherer, K. R.: Vocal expression of affect, The new handbook of methods in nonverbal behavior research, pp. 65-135 (2005).
- [6] 成瀬加菜 他: リアルタイムな変換聴覚フィードバックによる緊張緩和効果の基礎的検討, 研究報告エンタテインメントコンピューティング (EC), 2018-EC-48, Vol.17, pp.1-8 (2018).
- [7] Spielberger, C. D. et al.: Manual for the state-trait anxiety inventory (1970).
- [8] Brooks, A. W.: Get excited: Reappraising pre-performance anxiety as excitement, Journal of Experimental Psychology: General, Vol. 143, No. 3, pp. 11-44 (2014).
- [9] 牛山美和 他: 皮膚電位反応を用いた情動反応評価に関する基礎的検討, 紀要, Vol. 22, pp. 105-112 (1996).