

VR 模擬面接システム利用時における脈拍変動を用いた緊張感の評価:

VR 面接官の発話速度と声の高さによる影響の検討

Evaluation of anxiety using pulse rate variability associated with speaking rate and pitch of the VR-simulated job interviews

清水 翔太郎[†]
Shotaro Shimizuジャスティン・アンドリュー・リャオ[†]
Justin Andrew Liao神長 伸幸[†]
Nobuyuki Jincho菊池 英明[†]
Hideaki Kikuchi

1. はじめに

面接は就職活動や入学試験といった場面において、重要な観点の一つとなっている。一方、面接に苦手意識を持つ人は多く存在する。面接への苦手意識を払拭するためには、面接官役を他者に依頼して模擬面接を行うことが有効と考えられる。しかし、この練習方法は面接官役への依頼に加え、互いのスケジュールの調整も必要となるため、大きな労力がかかる。さらに、実際の面接官は基本的に初対面であるのに対して、面接練習の面接官役には身近な人物が選ばれることが多い。そのため、本番の面接の緊張感を再現することは困難である。

これらの問題を解決するために、マイナビの模擬面接シミュレーター[1]のような、一人でも模擬面接を行えるシステムが Web サービスとして提供されている。しかし、このような模擬面接システムは、ユーザーと面接官が PC 画面によって隔てられているため、現実の面接の緊張感を十分に再現できていないと考えられる。

また、より現実の面接に近い緊張感の実現が期待できる技術として、ヘッドマウントディスプレイ (以下 HMD) を用いたバーチャルリアリティ (以下 VR) 技術が存在する。Villani らは、VR 上の模擬面接によって現実の模擬面接を超える存在感と緊張感が得られる可能性を示した[2]。しかし、この研究は高い緊張感を得られるような VR 模擬面接環境の構築と評価を主題としており、就職面接への苦手意識を払拭できる機能については扱っていない。

就職面接への苦手意識を払拭するための機能としては、ユーザー自身によって模擬面接の緊張感を調整可能にすることで、緊張感の高い模擬面接に段階を踏んで挑戦できるようにする機能が考えられる。

ユーザーの緊張感に影響しそうな要素として、システムの面接官役 (アバター) が発する音声の韻律情報がある。韻律情報の操作によってユーザーの緊張感が調整できれば、緊張感の高い模擬面接に段階を踏んで挑戦できるようにする機能の提供に繋がると考えられる。

そこで、本研究では、VR を用いた模擬面接システムを開発し、特に面接官役 (アバター) が発する音声において、韻律情報がユーザーの緊張感に与える影響を検討する。また、開発した VR 模擬面接システムと現実の模擬面接の緊張感を比較するために、現実の模擬面接との緊張感の差異も検討する。

2. 本研究で扱う韻律情報

本研究では、音声の韻律情報のうち、 f_0 平均値と発話速度に注目する。

f_0 は、声の高さの主観的な感覚に対応する物理量である。川波らは、丁寧や喜びの感情を表す音声では f_0 平均値は上

昇し、ぞんざいや悲しみの感情を表す音声では f_0 平均値は低下する傾向にあることを示した[3]。この研究から、聞き手が話し手の感情を読み取る際には f_0 平均値が用いられていると考えられる。そして、聞き手が話し手の苛立ちや無関心といった感情表現を読み取った場合、聞き手の緊張感が高まることが予想される。そこで、話し手の音声の f_0 平均値を操作することで、聞き手による話し手の感情の評価に変化が生じ、得られる緊張感も変化するという仮説を立てた。本研究で使用する音声合成システムによって操作可能な pitch パラメータと f_0 平均値は、主観的な声の高さに影響を与える概念として共通している。

内田は、話し手の性格印象は音声の発話速度に影響を受けることを示した [4]。この研究から、聞き手が話し手の性格を推定する際には、発話速度が用いられていると考えられる。そして、聞き手が話し手の性格に対して苦手意識を持った場合、聞き手の緊張感が高まることが予想される。そこで、話し手の発話速度を操作することで、聞き手による話し手の性格印象の評価に変化が生じ、得られる緊張感も変化するという仮説を立てた。なお、先行研究における発話速度は PICOLA と呼ばれる話速変換技術を用いて操作されている。この操作は、主観的な発話速度を操作している点で、本研究における発話速度の操作と共通している。

3. 提案システム

韻律情報の操作が可能な VR 模擬面接システムを提案するために、システムの開発手法、機能および対話例について検討した。

3.1 開発環境

VR 空間の構築には統合型のゲーム開発環境である Unity を使用した。HMD には OculusRiftCV1 を使用した。面接官役として用いるアバターは、Web 上に公開されている 3D モデルを改変して作成した。擬人化エージェントの発話音声の作成には VoiceText Web API (HOYA 社製) を使用した。この API は、pitch と speed というパラメータで声の高さと発話速度をそれぞれ操作できる。

3.2 面接環境のデザイン

提案システムを起動し、HMD を装着すると、ユーザーは面接会場の中で面接官役のアバターと向かい合っている状態となる。ユーザー視点の面接環境を図 1 に示す。図 1 の状態から模擬面接が開始される。

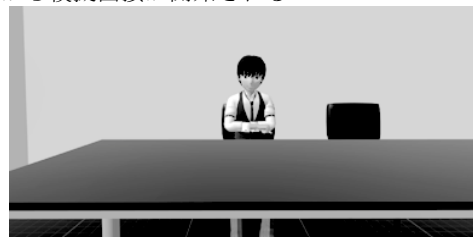


図 1 ユーザー視点の面接環境

[†] 早稲田大学, Waseda University.

3.3 対話例

対話は一問一答式で行われ、アバターは予め設定されていた内容のみを発話することができる。提案システムを利用した際の対話例は以下のようになる。

システム発話「本日はお忙しい中、弊社の面接にお越しいただき、ありがとうございます。今回面接を担当させていただきます。〇〇(人名)です。よろしくお願いします。」

ユーザー発話「よろしくお願いします。」

システム発話「それでは、早速ですが自己紹介を1分くらいでお願いします。」

ユーザー発話「はい。(自己紹介の内容)です。」

システム発話「ありがとうございました。それでは次に、あなたの長所を2分くらいで教えてください。」

ユーザー発話「はい。私の長所は(長所の内容)です。」

システム発話「ありがとうございます。それでは次の質問になりますが、あなたの短所を2分くらいで教えてください。」

ユーザー発話「はい。私の短所は(短所の内容)です。」

システム発話「ありがとうございました。それでは、次で最後の質問になりますが、あなたの人柄について周りの人からどういう人だと言われますか。2分くらいで教えてください。」

ユーザー発話「はい。私は周りの人から(人柄の内容)とよく言われます。」

システム発話「ありがとうございました。本日の面接は以上となります。お気をつけてお帰りください。」

4. 韻律情報が聞き手の緊張感に与える影響の評価

韻律情報が聞き手の緊張感に与える影響について評価するために、合成音声の自然性評定実験と、合成音声の緊張感評定実験を行った。

4.1 合成音声の自然性評定実験

VoiceText Web API を用いて作成可能な合成音声のうち、面接官役の発話として不自然な音声を除外するために、合成音声の自然性評定実験を行った。

4.1.1 参加者

日本語を母語とする大学生および大学院生 10 名(男性 5 名、女性 5 名、20 - 27 歳)が参加した。参加者には聴覚障害およびその経験がないことを確認した。また、参加に際して、研究の目的および方法を説明し、実験前に参加同意を得た。

4.1.2 刺激

VoiceText Web API は、声の高さについては 50% から 200%、発話速度については 50% から 400% という倍率を指定することができる。今回は、音声刺激として、それぞれ変更可能な範囲内において 10% 刻みで合成音声を作成し、音声刺激として用いた。そのため、声の高さを変更した際の自然性評定実験の音声刺激の数は 16 通り、発話速度を変更した際の自然性評定実験の音声刺激の数は 36 通りとなった。音声の発話内容は、「好きな食べ物は(食べ物の名前)です」とした。好きな食べ物の部分には、4 モーラかつ馴染みのある食べ物の名称(グラタン、唐揚げなど)がランダムに割り当てられた。

4.1.3 手続き

実験は、インターネット上にアンケートページで回答する形式で行った。アンケートページでは、音声の聴取にへ

ッドホンを使用することと、自然性の評価基準として合成音声の「声の高さ(ピッチ)」と日常生活で耳にする音声の「声の高さ(ピッチ)」との間に差異がないと感じた場合には自然な音声と評価し、かけ離れていると感じた場合には不自然な音声と評価することを教示した。参加者は、ランダムな順序で提示される音声刺激を聴取し、その自然性について回答を行った。自然性の評価には、「不自然(0 点) - 自然(6 点)」の両極尺度を用いた。

4.1.4 結果と考察

声の高さと自然性の関係を図 2 に示す。図 2 から声の高さが 120% のとき、自然性の平均点は最大となり、そこから声が高くなったり、低くなったりするにつれて、自然性の平均点は低くなる傾向があった。

同様に、発話速度と自然性の関係を図 3 に示す。図 3 から発話速度が 120% のとき、自然性の平均点は最大となり、そこから発話速度が速くなったり、遅くなったりするにつれて、自然性の平均点は低くなる傾向があった。

図 2 と図 3 から、声の高さと発話速度の自然性の平均点の分布は、正規分布に近いことが読み取れた。さらに、声の高さと発話速度がそれぞれ 100% のときの自然性を基準に 99% 信頼区間を取り、その下限を上回る音声を自然な音声と定義した。その結果、声の高さについては 90 - 130% の範囲が自然な音声となり、発話速度については、80 - 130% の音声が自然な音声となった。

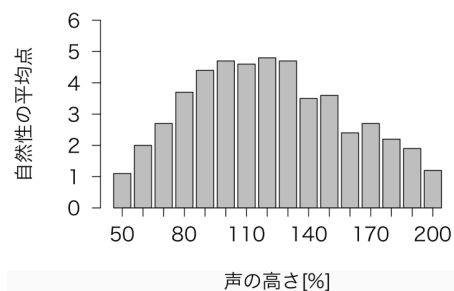


図 2 声の高さと自然性の関係

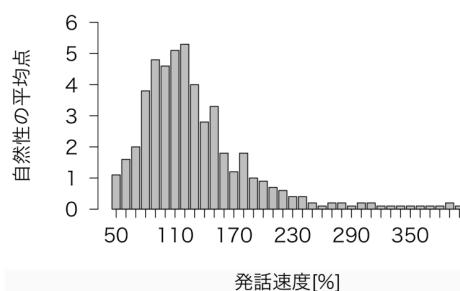


図 3 発話速度と自然性の関係

4.2 合成音声の緊張感評定実験

4.1 の実験において自然な音声と定義した合成音声をを用いて、音声の韻律情報と聞き手の緊張感の関連性を検討するため、合成音声の緊張感評定実験を行った。

4.2.1 参加者

日本語を母語とする大学生および大学院生 10 名(男性 5 名、女性 5 名、20 - 27 歳)が参加した。参加者には聴覚障害およびその経験がないことを確認した。また、参加に際して、研究の目的および方法を説明し、実験前に参加同意

を得た。なお、参加者には 4.1 の実験の参加者が 6 名 (男性 3 名, 女性 3 名) 含まれていた。

4.2.2 刺激

4.1 の実験において自然な音声と定義した, 声の高さが 90 - 130% の音声と発話速度が 80 - 130% の音声について, 交互作用を考慮するために, 10% 間隔で全ての組み合わせ (30 通り) を作成し, 音声刺激として用いた。音声の発話内容は, 「初めまして (人物の苗字) です。よろしくお願ひします。」とした。人物の名字の部分には, 3 モーラかつ馴染みのある名字 (加藤, 田中など) がランダムに割り当てられた。

4.2.3 手続き

大まかな手順は, 4.1.3 と同様である。聞き手の緊張感を高める音声の評価基準として, 「適切な受け答えができるだろうか」といった不安やあせりが高まったと感じる場合には, 「緊張する」と評価することを教示した。また, 音声の評価には「緊張しない (0 点) - 緊張する (6 点)」の両極尺度を用いた。

4.2.4 結果と考察

声の高さと緊張感の関係を図 4 に示す。図 4 から, 声の高さと聞き手の緊張感には負の相関が存在すると推測される。発話速度と緊張感の関係を図 5 に示す。図 5 から, 発話速度と聞き手の緊張感には弱い正の相関が存在すると推測される。また, 最も聞き手の緊張感を高めると評定された音声は声の高さが 90% かつ発話速度が 110% の音声であり, 最も聞き手の緊張を高めないと評定された音声は, 声の高さが 130% かつ発話速度が 80% の音声であった。なお, この二つの韻律情報の音声を, 「5. 提案システムの評価」において, 面接官役の 3D モデルの発話として使用する音声に使用した。

声の高さと発話速度がそれぞれ聞き手の緊張感に影響を与えた程度を考察するために, 緊張感の平均点を目的変数, 声の高さと発話速度を説明変数として, 以下の重回帰分析を行った。

$$(\text{緊張感の平均点}) = \beta_1 + \beta_2 * (\text{声の高さ}) + \beta_3 * (\text{発話速度})$$

重回帰分析の結果, 声の高さから緊張感の平均点への偏回帰係数 (β_2) は, 5% 水準で負の方向に有意であった ($\beta_2 = -2.450$, $t = -6.029$, $p < 0.01$)。このことから, 声が低くなるほど聞き手の緊張感が高くなることが読み取れた。

発話速度から緊張感の平均点への偏回帰係数 (β_3) は, 5% 水準で正の方向に有意傾向にあった ($\beta_3 = 0.645$, $t = 1.919$, $p = 0.066$)。このことから, 発話速度が速くなるほど聞き手の緊張感が高くなる傾向が読み取れた。

なお, 今回の重回帰分析は, 4.1 の実験において予め自然な音声とされた音声データのみを対象としているため, 回帰式も自然な声の高さや発話速度の範囲内でのみ適用することに注意する必要がある。

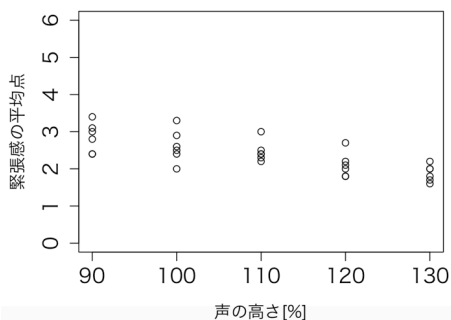


図 4 声の高さと緊張感の関係

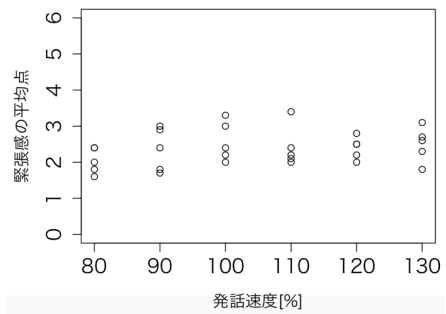


図 5 発話速度と緊張感の関係

5. 提案システムの評価

開発したシステムの面接官役 (アバター) が発する音声について, 韻律情報を操作することで緊張感が実際に変化するかを検証するために, 4.2.4 で得られた二種類の音声を使用した VR 模擬面接の緊張感を比較した。また, 提案システムと現実の模擬面接におけるユーザーの緊張感の差異を検討するために, 現実の模擬面接での緊張感との比較も行った。緊張感の評価指標には, 脈拍数と状態特性不安検査 (STAI) の状態不安得点を使用した。脈拍数は不整脈や血液量の減少などによって脈拍が観測できない場合を除き, 心拍数に等しい。心拍数は, 不安や緊張によって上昇することが知られている [5]。状態不安は, 不安を喚起する事象に対する一過性の状況反応であり, 脅威であると知覚された場面では状態不安得点は高くなるが, 危険性が全くないかほとんどない場面では状態不安得点は低くなる [6]。

5.1 参加者

日本語を母語とする大学生 33 名 (男性 20 名, 女性 13 名) が参加した。参加者は, 新卒採用としての就職面接を未経験で, 4.1 および 4.2 の実験に参加していない者に限定した。なお, 入学試験やアルバイト, インターンシップなどの面接経験については不問とした。

5.2 手続き

33 名の参加者のうち, 現実の模擬面接に 11 名 (男性 7 名, 女性 4 名), 4.2 の実験において最も聞き手の緊張感を高めるとされた音声を用いた VR 模擬面接 (以下, 緊張する音声を用いた VR 模擬面接と表記) に 11 名 (男性 6 名, 女性 5 名, 4.2 の実験において最も聞き手の緊張感を高めないとされた音声を用いた VR 模擬面接 (以下, 緊張しない音声を用いた VR 模擬面接と表記) に 11 名 (男性 7 名, 女性 4 名) 参加してもらった。

実験前, 参加者には腕時計型の脈拍計を装着してもらい, 安静時の脈拍を 5 分間測定した。その後, 実験の注意事項として, 面接官に質問をしないことと, 回答時間の指定をそれほど厳密に守る必要はないことを教示した。

面接官の発話内容は, 3.3 のシステム発話例の通りとした。現実の模擬面接の面接官役は, 参加者と面識のない男子大学生一名に依頼した。

VR 模擬面接の進行に際しては, システムではなく実験者が参加者の発話終了を判定し, 手動で次のシステム発話を開始させる方法 (Wizard of Oz 法) を採用した。実験後は, STAI の質問紙を用いて, 状態不安得点と特性不安得点を測定した。最後に, 脈拍を 5 分間測定した。なお, 脈拍の測定は安静時から実験後まで継続された。脈拍変動を

検討する際は、実験の進捗から実験時間を五つのフェーズ（実験前安静時、注意事項説明時、模擬面接時、STAI 回答時、実験後安静時）に分割した。

5.3 結果と考察

各条件・各フェーズにおける平均 bpm を図 6 に示す。図 6 から、全ての条件において模擬面接時までは脈拍が上昇し、模擬面接後は脈拍が下降する傾向があった。また、面接前安静時の時点で既に平均 bpm が異なることから、脈拍数は個人差が大きいことも読み取れた。脈拍数の個人差が大きいことから、VR 模擬面接における面接時の脈拍数が、現実の模擬面接における面接時の脈拍数を上回っていたとしても、VR 模擬面接の緊張感が現実の模擬面接の緊張感を超えているとは言い切れない。また、各条件の特性不安得点と状態不安得点からは、各条件間の明らかな差異は読み取れなかったが、特性不安が高い参加者は状態不安も高くなる傾向が確認された。

模擬面接の条件およびフェーズによる平均 bpm の差異を考察するために、bpm を目的変数、模擬面接の条件（現実、緊張する音声を用いた VR、緊張しない音声を用いた VR）、フェーズ（模擬面接時、実験後安静時）、STAI の特性不安得点を説明変数、参加者ごとの bpm を個体差とする線形混合モデルを用いて以下の分析を行った。なお、特性不安とは脅威を与える状況に対する反応の傾向であり、不安傾向の個人差を示す指標である[6]。

$$(\text{平均 bpm}) = \beta_1 + \beta_2 * (\text{模擬面接の条件}) * (\text{フェーズ}) + \beta_3 * (\text{特性不安}) + (\text{脈拍の個人差})$$

統計モデリングの結果の要約統計量を表 1 に示す。なお、表 1 は緊張しない音声を用いた VR 模擬面接における模擬面接時の平均 bpm を基準としている。表 1 から、面接後安静時は模擬面接時より低い bpm を示すことが 5%水準で有意となった。このことから、いずれの条件においても模擬面接時の参加者は模擬面接後安静時より緊張感が高まっていたと考えられる。また、現実の模擬面接と緊張する音声を用いた VR 模擬面接は緊張しない音声を用いた VR 模擬面接と比較して高い平均 bpm を示すことが 5%水準で有意となった。このことから、声の高さや発話速度の変化が面接時の緊張感に影響することが示唆された。そして、現実の模擬面接は緊張しない音声を用いた VR 模擬面接と比較して、面接終了後に大きく平均 bpm が下降することが有意傾向となった。このことから、現実の模擬面接の参加者は緊張しない音声を用いた VR 模擬面接の参加者より実験後の緊張感が低下する傾向にあったと考えられる。

さらに、模擬面接の条件から状態不安得点を考察するために、状態不安得点を目的変数、模擬面接の条件と特性不安得点を説明変数として以下の重回帰分析を行った。

$$(\text{状態不安得点}) = \beta_1 + \beta_2 * (\text{模擬面接の条件}) + \beta_3 * (\text{特性不安得点})$$

重回帰分析の結果、それぞれの条件間において、状態不安得点に有意な差は確認されなかった。

表 1 各条件・各フェーズと平均 bpm の関係

	偏回帰係数	標準誤差	t値	p値
条件 緊張する音声を用いたVR	8.124	3.357	2.420	0.022
条件 現実	6.269	3.127	2.059	0.048
フェーズ 実験後安静時	-7.091	2.005	-3.536	0.001
特性不安	0.003	0.136	0.024	0.981
条件 緊張する音声を用いたVR * フェーズ 実験後安静時	-0.909	2.836	-0.321	0.751
条件 現実 * フェーズ 実験後安静時	-5.636	2.836	-1.987	0.056

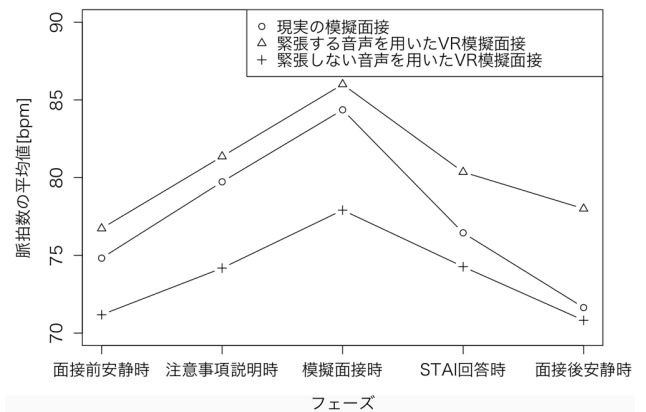


図 6 各条件・各フェーズにおける平均 bpm

6. まとめ

本研究では、VR を用いた模擬面接システムを開発し、特に面接官役（アバター）が発する音声の韻律情報の変化がユーザーの緊張感に与える影響を検討した。そして、現実の模擬面接と緊張に影響する韻律情報を操作した二つの VR 模擬面接の三条件について、緊張感の差異を検討した。その結果、いずれの模擬面接についてもユーザーは面接時に脈拍数が上昇すること、韻律情報の操作によって脈拍数が上昇することが確認された。一方、STAI の状態不安得点は、三条件間で有意な差が確認されなかった。今後は、脈拍数と状態不安得点の一方のみ有意差が確認された理由について考察し、より信頼性の高い実験計画および分析手法を検討したい。

また、韻律情報の操作による脈拍数の変化が確認されたことから韻律情報の操作によってユーザーの緊張感を調整できることが示唆されたため、今後はユーザーが緊張感の高い模擬面接に段階を踏んで挑戦できるようなシステムおよびユーザーインターフェースの実装に注力する。

さらに、今後はユーザーの面接スキル向上という視点からのシステム評価に加え、VR 模擬面接によって得られる緊張感をさらに高める手法の検討に取り組んでいきたい。

参考文献

- [1] マイナビ, 模擬面接シミュレーター - 就活支援 - マイナビ 2017, (online), <http://job.mynavi.jp/conts/2017/mensetsu/mogimen/sci/mogi_01.html> (最終閲覧日: 2016 年 12 月 23 日).
- [2] D. Villani, C. Reppeto, P. Cipresso, G. Riva, "May I experience more presence in doing the same thing in virtual reality than in reality? An answer from a simulated job interview", *Interacting with Computers*, Vol.24, pp.265-272 (2012).
- [3] 川波弘道, 広瀬啓吉, "態度・感情音声における韻律的特徴の考察", 電子情報通信学会技術研究報告, SP, 音声 Vol.97, No.396, pp.73-80(1997).
- [4] 内田照久, "音声の発話速度が話者の性格印象に与える影響", *心理学研究*, Vol.73, No.2, pp.131-139(2002).
- [5] 大須賀美恵子, "心的状態の指標としての心拍・心拍変動", *ヒューマンインターフェース学会誌*, Vol.6, No.1, pp.9-14(2004).
- [6] 肥田野直, 福原真知子, 岩脇三良, 曾我祥子, Charles D. Spielberger, "新版 STAI マニュアル", p.5(2000).