

対話において上手く褒めるための  
頭部・顔部の振舞いの分析

令和3年度 修士論文

日本大学大学院 総合基礎科学研究科  
地球情報数理科学専攻 宮田研究室

6120M10 大西 俊輝

# 概要

日常生活や社会活動において、褒める行為は重要なコミュニケーションのひとつである。褒める行為は、対象の行動や性格に向けられた称賛を表現する言語的・非言語的な行動であると考えられている。褒めることで良好な人間関係の構築やパフォーマンスの向上が期待されているが、どのような褒め方をするとよいか悩む人も少なくない。加えて、心理学や教育学の分野において褒める行為による効果や影響が明らかにされているが、相手を上手く褒めるための言語・非言語行動は明らかにされていない。そこで、実際の対話データを用いて、相手を上手く褒めるための言語・非言語行動を明らかにする取り組みを行う。本研究では、対話における褒め方の上手さと人間の行動の関係を分析する取り組みとして、言語・非言語行動の中から頭部と顔部の振舞いに着目し、相手を上手く褒めるためにはどのような頭部と顔部の振舞いが重要であるか明らかにする取り組みを行う。はじめに、頭部と顔部の振舞いと、褒め方の上手さの評価値を含む対話コーパスを作成した。次に、頭部と顔部の振舞いに関連する特徴量を用いて褒め方の上手さの評価値を推定する機械学習モデルを構築し、どのような頭部と顔部の振る舞いが重要であるかを検証した。その結果、視線の方向に関する振る舞い、口角や口の開きに関する振る舞いが相手を上手く褒めるためには重要であることが明らかになった。その結果、褒める人の視線の方向に関する振る舞い、口角や口の開きに関する振る舞いが相手を上手く褒めるためには重要であることが明らかになった。

本稿の貢献は、話者（褒める人と褒められる人）の頭部と顔部の振舞いに着目し、相手を上手く褒めるためにはどのような振舞いが重要であるか明らかにしたことである。

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	研究の背景	2
1.2	研究の目的	2
1.3	本論文の構成	3
<b>第2章</b>	<b>関連研究</b>	<b>4</b>
2.1	褒める行為に関連する研究事例	5
2.2	個人の性格や能力を推定する研究事例	6
2.2.1	個人の性格特性を推定する研究事例	6
2.2.2	個人の能力を推定する研究事例	8
<b>第3章</b>	<b>研究課題</b>	<b>13</b>
3.1	問題の定義	14
3.2	研究課題の設定	15
<b>第4章</b>	<b>対話コーパス</b>	<b>17</b>
4.1	2者対話の収録	18
4.2	アノテーション	19
<b>第5章</b>	<b>上手く褒めるための頭部・顔部の振舞いの分析</b>	<b>22</b>
5.1	特徴量抽出	23
5.2	褒め方の上手さの推定モデルの構築	24
5.3	褒め方の上手さの推定モデルの結果	25
5.4	上手く褒めるための頭部・顔部の振舞いの分析	26
5.4.1	褒める人の頭部・顔部の振舞いの分析	27
5.4.2	褒められる人の頭部・顔部の振舞いの分析	29
5.5	本分析で得られた知見	30
<b>第6章</b>	<b>結論</b>	<b>32</b>
	参考文献	35
	研究業績	40

# 目 次

4.1	2者対話の様子	18
4.2	Praise スコアの分布 (N=228)	20
5.1	頭部の x 軸, y 軸, z 軸方向の座標系	24
5.2	特徴量を抽出する範囲	24
5.3	Model P における特徴量の重要度上位 10 件	27
5.4	Model P+R における特徴量の重要度上位 10 件	28
5.5	実際の対話シーン (Praise シーン) の画像と特徴量の値の一例	28

# 表 目 次

2.1	本研究が属する研究領域 . . . . .	5
4.1	発話シーンと Praise シーンの情報 . . . . .	19
5.1	Action Units の内容 . . . . .	25
5.2	各指標の平均値 (N=100) . . . . .	26
5.3	Praise スコア高群と低群における重要度の高い特徴量の中央値 . . . . .	29
5.4	重要度が高い特徴量と Praise スコアの相関係数 . . . . .	30

# 第1章 序論

## 1.1 研究の背景

日常生活や社会活動において、褒める行為は重要なコミュニケーションのひとつである。初対面の相手と話す場合、相手の見た目やエピソードを積極的に褒めることで相手が心を開き、話が弾み、その後人間関係が円滑になることが期待される。あるいは、同僚や部下の行為を適切に褒めることで彼らのパフォーマンスを最大化することが、組織マネジメントの観点からも重要であるとの認識が広がっている。褒める行為は、対象の行動や性格に向けられた称賛を表現する言語的・非言語的な行動であると考えられている [1][2]。さらに、褒める行為は、褒める人から褒められる人への一方的な意思の伝達ではなく、複雑な社会的コミュニケーションであり、褒められる人の役割は褒める人の役割と同じくらい重要であるされている [3]。

褒めることで良好な人間関係の構築や、褒めた相手のパフォーマンスの向上が期待されているが、どのような褒め方をすればよいか悩む人も少なくない。褒める行為に関する研究事例も一定数存在するが、その多くは、特定のタスク後に実施するアンケートや主観評価の結果を用いて分析を行い、褒める行為の効果や影響を調査するにとどまっている。すなわち、相手を上手く褒めるために言語・非言語行動をどのように用いるとよいかは明らかにされていない。

## 1.2 研究の目的

1.1節で述べた研究の背景より、相手を上手く褒めるために言語・非言語行動をどのように用いるとよいか明らかにする必要がある。そこで、人間の言語・非言語行動を利用して、対話における褒め方の上手さを分析する取り組みを行う。具体的には、対話における褒め方の上手さと人間の行動の関係を明らかにするために、話者（褒める人、褒められる人）の頭部と顔部の振舞いに着目し、相手を上手く褒めるためにはどのような行動が重要であるのか調査・分析する取り組みを行う。はじめに、頭部と顔部の振舞いと、褒め方の上手さの評価値を含む対話コーパスを作成する。次に、頭部と顔部の振舞いに関連する特徴量を用いて褒め方の上手さの評価値を推定する機械学習モデルを構築し、どのような頭部と顔部の振舞いが重要であるかを検証する。

さらに、本研究では、先行研究 [4][5][6][7] の改善を行う。具体的には、対話データの拡張、発話シーンの再定義、アノテータの人数やアノテーション方法の変更、機械学習モデルの構築方法の変更を行う。加えて、上手く褒めるために重要と判断した振舞いが、褒め方の上手さの評価値とどのような関連があるのか明らかにする。

上記をふまえ、本研究では褒める人と褒められる人の頭部、顔部の振舞いに着目し、相手を上手く褒めるためにはどのような振舞いが重要であるのか明らかにすることを研究目的とする。これにより、対話において相手を上手く褒めるためにはどのような行動が重要であるか明らかになり、相手を上手く褒めるための能力の向上や褒め方の上手さを評価するシステムの開発が期待される。

### 1.3 本論文の構成

本論文の構成は次のとおりである。

2章では、褒める行為に関連する研究事例と個人の性格や能力を推定する研究事例について述べる。

3章では、既存研究における問題を定義し、先行研究の改善点をふまえた上で本論文における研究課題について述べる。

4章では、2者対話の収録、アノテーション、褒め方の上手さの評価について述べる。

5章では、褒め方の上手さを推定する機械学習モデルの構築と上手く褒めるための頭部・顔部の振舞いの分析について述べる。

最後に6章にて、本論文の結論を述べる。

## 第2章 関連研究

本研究では、対話における話者の非言語行動を用いて、相手を上手く褒めるためにはどのような行動が重要であるのか明らかにする取り組みを行う。本章では、この取り組みを行うにあたり、褒める行為に関連する研究事例と個人の性格や能力を推定する研究事例について述べる。本研究が属する研究領域を表2.1に示す。2.1節では、褒める行為に関連する研究事例について紹介する。これらは、褒める行為に着目しているという点で本研究と関連している。2.2節では、個人の性格や能力を推定する研究事例について紹介する。これらは、人間の言語的・非言語的行動を利用して、特定のタスクや対話中の行動・能力を分析するという点で本研究と関連している。

表 2.1: 本研究が属する研究領域

研究分野	研究事例
教育学, 心理学	<ul style="list-style-type: none"> <li>・褒める行為の機能を明らかにする研究</li> <li>・褒める行為の効果を明らかにする研究</li> </ul>
マルチモーダルインタラクション (社会的信号処理)	<ul style="list-style-type: none"> <li>・人間の言語・非言語行動を利用して個人の性格特性を推定する研究</li> <li>・人間の言語・非言語行動を利用して個人のスキルやパフォーマンスを推定する研究</li> </ul>

## 2.1 褒める行為に関連する研究事例

本節では、褒める行為の機能や効果に関連する研究事例について述べる。多くの研究事例では、褒める行為の機能や褒めることによる効果などを明らかにしている。ただし、大半の研究事例は教育シーンにおける褒める行為を扱っている。日常生活やビジネスシーンにおける対話中の褒める行為を扱う研究事例は少なく、心理カウンセラー、企業の管理職経験者、スポーツの指導経験者が自らの経験を書籍にまとめている事例が多い。

はじめに、褒める行為の機能に関する研究事例について述べる [8][3][9]。Brophy[8]は、教師による褒める行為は、生徒に対する強化の手段として広く推奨されており、生徒の振る舞い・学業成績の強化以外に様々な機能があることを示している。教師による褒める行為は、(1) 自然発生的な驚嘆・称賛の表現、(2) 過去に行った指摘に対するバランスの維持、(3) 代理的な強化（ある生徒を褒めることで、そうすべきと残りの生徒に伝える）、(4) 否定表現の回避（否定表現をしないで褒めることで指導する）、(5) 緊張をほぐす手段、(6) 生徒からの誘発、(7) 移行の儀式（タスクが終了し、次のタスクに移行する合図）、(8) 慰めの賞・励ましの8通りに分類できると考えられている。加えて、生徒や教師間の親密な関係を築く手助けになる望ましい行為であると考えられている。

Henderlongら [3]は、子どもの内発的な意欲や忍耐力に対する褒める行為の効果は、価値のあるものから価値がわずかなもの、不利益なものまで、複雑かつ多様であると述べて

いる。また、褒める行為は、年齢、性別、文化などの受け手の特性によって、受け手のモチベーションに異なる影響を与える可能性があることを述べている。

Ennisら[9]は、褒める行為の効果を高めるためには、行動の成果に基づいて褒めるべきである、褒める対象の行動を明示すべきである、誠実に褒めるべきであると述べている。

次に、褒める行為の効果・影響を明らかにする研究事例について述べる[10][11]。Andersonら[10]は、4～5歳児を対象に絵画を描くという課題を用いて実験を行い、褒める行為の効果を一明らかにしている。参加者に「絵を描いてほしい、もし何か絵を描いてくれたらご褒美（報酬）をあげる」と伝えた上で8分間絵を描いてもらい、描画時間を比較している。報酬の種類は、よくできました賞のようなシンボル、お金、ポジティブな言語的報酬（例：その絵は本当にいいね）であった。報酬を与えずに行った実験での描画時間と、報酬を与えて行った実験での描画時間を比較すると、ポジティブな言語的報酬を与えられた場合は絵を描く時間が長くなり、その他の場合は絵を描く時間が短くなることを明らかにしている。

Stipekら[11]は、褒める行為が子供の自尊感情に与える影響について子供の感情表出に着目して明らかにしている。2～5歳児を対象にパズル課題に取り組んでもらい、課題に成功した後、「本当に上手にパズルができたね」と成功したことについて褒めるもしくは「パズルが終わったね」と成功した事実についてコメントするという実験を行っている。コメントされた後の子供の表情・身振りなどを確認した結果、「本当に上手にパズルができたね」と成功したことについて褒められた子供は、成功した事実についてコメントされた子供よりも、笑顔や堂々としたポーズなどのポジティブな表出を多くすることを明らかにしている。

## 2.2 個人の性格や能力を推定する研究事例

本節では、個人の性格や能力を推定する研究事例について述べる。2.2.1項では、人間の言語・非言語行動を利用して、個人の性格特性を推定する研究事例を紹介する。2.2.2項では、人間の言語・非言語行動を利用して、個人のプレゼンテーション、コミュニケーション、リーダーシップ、共感、自己開示のスキルを推定する研究事例について紹介する。

### 2.2.1 個人の性格特性を推定する研究事例

本項では、人間の言語・非言語行動を利用して、個人の性格特性を推定する研究事例について述べる[12][13][14]。これらの研究事例では、個人の性格特性を推定するために、人間の言語・非言語行動を記録したコーパスを利用もしくは作成し、コーパスから抽出した特徴量を用いて機械学習モデルを構築している。

Batrincaら[12]は、human-machine interaction (HMI) と human-human interaction (HHI) の2つのシナリオで収集した音声・映像データから、Big Fiveの性格特性（外向性、誠実性、情緒不安定性、開放性、調和性）を自動認識している。HMIとHHIという

2つの異なるシナリオにおける性格特性の現れ方を分析するために、マルチモーダルデータセットを収集している。このデータセットは、Andersonら[15]によって導入されたマップタスクコーパスを改良したものである。マップタスクは、心理学、社会学、言語学の研究で広く使われている協力的なタスクであり、タスクによって引き出された自発的な対話やインタラクションを収集することができる。1つ目のシナリオでは、実験参加者はコンピュータ（HMI）と対話を行うが、実験者がシステムの応答を操作してデータを収集している。一方で、2つ目のシナリオ（HHI）では、実験参加者は他の人と直接対話することを明示してデータを収集している。2つのシナリオのデータセットから視覚、音声（アノテーションの抽出による音声と自動抽出による音声）に関する特徴量を抽出している。これらの特徴量から分類モデルを構築し、性格特性を推定している。その結果、1つ目のシナリオでは正解率が0.67、2つ目のシナリオでは正解率が0.70であり、性格特性の自動認識の実現可能性があることを示唆している。

Valenteら[13]は、自発的な会話における話者の性格特性を推測するためのアノテーションと検証を行っている。データセットは、AMIコーパス[16]を利用している。このコーパスは、会議室で撮影されたミーティングのコーパスであり、ミーティング参加者それぞれの音声と映像が記録されている。映像は、ミーティングの全体像と参加者をクローズアップした映像の両方が記録されている。加えて、コーパスにはシナリオに沿ったミーティングと非シナリオのミーティングが含まれており、シナリオミーティングでは4人の参加者がプロジェクトマネージャ、マーケティングエキスパート、ユーザインタフェースデザイナー、インダストリアルデザイナーからなるデザインチームとして新しいリモコンをデザインするシナリオとなっている。プロジェクトマネージャは、他の発言者と議論すべきいくつかの項目を含むアジェンダに従って会議を進行する。このコーパスは、様々な情報（役割、発言時間、単語、対話行為）が書き起こされている。上記のコーパスから、言語特徴量（単語のn-gram、対話行為）と非言語特徴量（韻律、発話の活性度、発話の重複、発話のポーズ）を抽出している。これらの特徴量から、ブースティングを用いて性格特性を推測する取り組みを行っている。その結果、外向性、誠実性、情緒不安定性は、それぞれ74.5%、67.6%、68.7%の精度で推測できることが明らかにされている。加えて、非言語特徴量は、言語特徴量より有用であることも明らかにしている。

Jayagopiら[14]は、対面での相互作用における発話と視線のパターンを特徴づける集団行動の手がかりを定義し、抽出するためのフレームワークを提案している。データセットは、ELEAコーパス（Emergent LEADER Corpus）[17]から18のグループインタラクションを用いている。ELEAコーパスは、ミーティングにおける個人の性格特性やリーダーシップに関する情報と40件（約10時間）のミーティングを映像と音声で記録しているコーパスである。上記のコーパスから、グループにおける非言語行動の抽出を行い、グループでの社会的・心理的構成（性格特性、対人認識、能力等）と関連づけることを行っている。非言語行動は、個人に関する特徴量（発話状態、発話長、発話ターン、割り込み成功、割り込み失敗、バックチャンネル、頭部姿勢、視線方向）、グループに関する特徴量（ミーティングの継続時間、発話回数、無音区間、発話の重複区間、発話の非重複区間）を抽出している。上記の特徴量を離散化するためにBag of nonverbal patterns（Bag-of-NVPs）を定

義し，LDA（Latent Dirichlet Allocation）トピックモデルの構築を行っている．その結果，グループに関する特徴量とトピックの両方が，すべての社会的・心理的構成と有意な相関を持つことが示唆されている．

### 2.2.2 個人の能力を推定する研究事例

本項では，人間の言語・非言語行動を利用して，個人の能力を推定する研究事例について述べる．個人の能力として，プレゼンテーション [18][19][20]，コミュニケーション [21][22]，リーダーシップ [23]，面接 [24]，共感 [25][26]，自己開示 [27] のスキルを推定する研究事例について述べる．これらの研究事例では，個人の能力を推定するために，人間の言語・非言語行動を記録したコーパスを利用もしくは作成し，コーパスから抽出した特徴量を用いて機械学習モデルを構築している．

Ramanarayanan ら [18] は，人間が評価したプレゼンテーション能力のスコアを予測する際の比較を行っている．データセットは，プレゼンテーション中の3D身体動作，音声，映像を記録した56個のプレゼンテーションから構成されている．上記のデータセットから，頭部方向，視線方向，表情，運動量や手の動きの程度を特徴量として抽出している．これらの特徴量を用いて，プレゼンテーション能力のスコアを予測するタスクを行ったところ，すべての特徴量を組み合わせた機械学習モデルにおいて最も良いスコアが予測されることが確認された．

Chen ら [19] は，プレゼンテーションを評価するためのマルチモーダルスコアリングモデルを構築している．データセットは，プレゼンテーション中の3D身体動作，音声，映像を記録した56個のプレゼンテーションから構成されている．上記のデータセットから，発話（動詞句（VP），T-units（T），節（C）など9種類の構文構造の頻度，平均節長（MLC），T-unitsあたりの動詞句（VP/T）など14種類の構文複雑度），発話内容（発話速度，韻律変化，ポーズ，発音），非言語行動（運動量，手の動き，頭の向き）に関する特徴量を抽出している．これらの特徴量を用いて，プレゼンテーションの評価を予測する回帰モデルを構築している．機械学習モデルは，Random Forests（RF）とSupport Vector Machine（SVM）の2種類を用いている．その結果，どちらのモデルにおいても発話，音声，非言語行動に関する特徴量をそれぞれ用いることで，それぞれのモダリティがプレゼンテーション全体のパフォーマンスを予測するために有用であることが示唆されている．加えて，3つのモダリティのうち最も有用であったのは音声特徴量であること，RFモデルでは発話と音声の両方の特徴を組み合わせることでパフォーマンスが向上することが確認された．

Yagi ら [20] は，言語・非言語行動からプレゼンテーションスキル推定するためのインスタンスの重み付けの適用法を提案している．はじめに，58のプレゼンテーションセッションで記録された参加者の音声，身体動作，テキストデータを含むマルチモーダルプレゼンテーションデータセットを構築している．このデータセットには，人事経験のある2人の外部評価者によって評価されたスコアも含まれている．プレゼンテーションスキルを推定するために，音声，身体動作，言語に関する特徴量を抽出している．次に，マルチ

モーダル・インスタンスの重み付け適応に基づいた回帰モデルに対して、入力値を結合するモデル (Early fusion) と出力値を結合するモデル (Late fusion) の2つのアプローチを提案している。実験の結果、インスタンス重み付け適応を用いた入力値を結合する回帰モデルは、プレゼンテーションの目標要素の明確さに対する回帰精度を示す相関係数が改善されることを明らかにしている。

Park ら [21] は、オンライン・ソーシャル・マルチメディア・コンテンツにおける話者の説得力を予測するために、複数のコミュニケーション・モダリティからの言語的行動と非言語的行動を利用したアプローチを提案している。データセットは、商品レビューのビデオを格納している Web サイトにおけるレビュー動画と商品の評価をコーパスとして用いている。各商品レビューには、特定の商品について話す話者の動画と、その商品に対する話者の直接評価が星1つ (最も否定的な評価) ~ 星5つ (最も肯定的な評価) の整数値が記録されている。その中から、合計 1000 本の映画レビュービデオを収集している。収集したレビュービデオの詳細は、肯定的なレビュー動画 500 本 (評価が5つ星)、否定的なレビュー動画 500 本 (1つ星が 216 本, 2つ星が 284 本) である。コーパスに含まれる各ビデオは、特定の映画について話す1人の人物を正面から撮影したものであり、ビデオの平均長は約 94 秒である。上記のデータセットから、音声 (ピッチ, フォルマント, 声質, MFCC (mel-frequency cepstral coefficients)), 言語 (発話率, ポーズ, ポーズフィルター, 発話の妨害率, 発話の吃り), 視覚 (感情, ポジティブ/ネガティブの値, Action Units[28], 視線移動, 姿勢) に関する特徴量を抽出している。これらの特徴量を用いて、話者の説得力を予測するタスクを行っている。その結果、3つのモダリティを融合して学習させた予測モデルは、それぞれのタイプのユニモーダル予測モデルと比較して統計的に有意に良い結果を示すことが確認されている。これより、話者の説得力の予測に言語・非言語行動が有用であることが示唆されている。

Sanchez-Cortes ら [23] は、話し方、韻律的特徴、視覚的行動、および動きといった非言語的行動を組み合わせて、グループにおける個人のリーダーシップを分析している。データセットには、ELEA コーパス [17] を用いている。上記のコーパスから話者交替、発話、頭部運動、身体運動に関する特徴量を抽出している。グループにおける個人のリーダーシップに関連する要素を特定するための話者交替、発話、頭部運動、身体運動に関する特徴量の有効性を評価した。相関分析の結果、グループにおけるリーダーシップのある人物は、積極的で支配的な人物であり、最もよく話し、話者交代やポーズが多く、声のトーンやエネルギーが長く変化することを明らかにしている。話者交替、発話、頭部運動、身体運動に関する特徴量を組み合わせることは、単一モダリティよりも優れた結果を示している。

Okada ら [22] は、人事管理の経験を持つ外部評価者によって評価された個人のコミュニケーションスキルを計算的に分析している。コミュニケーションスキルを推定するために、グループディスカッションが記録されている MATRICS コーパス [29] を用いている。MATRICS コーパスは、4名1組のグループが10組 (計40名) で構成されており、各組が3回のディスカッションを行っている。ディスカッションの内容は、対象とする人物などの順位付けに関するタスク、事前に情報が共有された企画タスク、事前に情報が共有されていない企画タスクの3つをランダムに行っている。MATRICS コーパスには、カメラ、

マイク、深度センサなどを用いて記録したマルチモーダルデータとコミュニケーションスキルに関するスコアデータが含まれている。上記のコーパスから言語行動（品詞、対話行為）、非言語行動（発話ターン、韻律情報、頭部行動）に関する特徴量を抽出した。これらの特徴量が、コミュニケーションスキルを推定するために有用であるのか分析している。具体的には、回帰モデルを構築し、タスクごとにコミュニケーションスキルを最もよく推定することができる特徴量の組み合わせを明らかにしている。その結果、参加者が事前に情報が共有されていないタスクにおいて言語行動に関する特徴量が重要であることが明らかになった。この結果より、話者はタスクに応じて異なる方法でコミュニケーションスキルを発揮していることが示唆された。

Nguyen ら [24] は、Hirability における応募者と面接官の非言語的な行動を用いて、面接での採用可能性を予測するためのフレームワークを提案している。Hirability とは、職種、面接の内容、面接の行われ方に依存する社会的構成の概念である [30]。データセットは、62 件（計 670 分）の採用面接をカメラとマイクで記録したコーパスを用いている。また、Hirability を評価する標準的な方法は存在していないことから、構造化された面接の 4 つの行動（コミュニケーション能力、説得力、誠実な仕事振り、ストレスへの耐性）に関する質問による 4 つのスコアと、面接の全過程に関連する 1 つのスコアからなる 5 つの Hirability スコアを定義している。上記のデータセットから、音声（発話のポーズ、発話時間、発話の流暢さ）、視覚的（頷き、頭部の動き、視線、表情、外見）、音声と視覚の総合的（音声のバックチャネル、視覚的バックチャネル、音声と視覚的のバックチャネル、相互発話、同時頷き）に関する特徴量を抽出している。これらのスコアと特徴量を用いて、応募者の採用可能性の予測を行っている。その結果、応募者の行動だけでなく、面接官の合図も Hirability スコアと相関があることがわかった。さらに、回帰タスクを行ったところ、非言語的な行動から Hirability スコアを予測することが可能であることが示唆されている。最後に、特徴量の予測妥当性を分析したところ、最も予測可能な特徴量は音声特徴量と面接官の視覚的特徴量であることが確認された。これより、面接官がより多くの視覚的なバックチャネルを表示することで、応募者の行動反応を生み出したことを示唆している。これは、面接官を見るだけで、面接の結果を予測できる可能性を示している。

Ishii ら [25] は、Davis の対人反応性指標を用いて測定した共感スキルレベルに応じて、話者交替/継続における視線行動と対話行為について分析している。はじめに、4 名 1 組のグループによるディスカッションをカメラ、マイクを用いて撮影を行い、発話、対話行為、視線行動、共感スキルレベルが含まれるマルチモーダルコーパスを構築している。次に、共感スキルが高い人と低い人では、5 つの対話行為のカテゴリ（提供、自己開示、感情移入、転回、他者）において発話に伴う話者交替/継続の視線行動が定量的に異なることを分析している。その結果、共感スキルの低い参加者は、対話行為のカテゴリが自己開示の場合、話者継続時に聞き手とのアイコンタクトを避ける傾向があった。一方で、対話行為のカテゴリが共感の場合、聞き手とのアイコンタクトを開始する傾向が見られた。共感スキルの高い参加者は、話し手の発話の対話行為のカテゴリが提供や共感の場合、話者交替時に話し手から目をそらすことが多い。一方で、話し手の発話の対話行為のカテゴリが展開の場合、共感スキルの違いによる視線の動きの違いは見られなかった。GTP と対話

行為の情報をを用いて、共感スキルレベルを推定するモデルを構築した。GTPは、任意の時間区間内で、現話者、非話者、人物以外の注視対象に対して、どのような順序で視線を向けたか、また人物を注視した際に視線交差が起きたかという時系列的な遷移をN-gramパターンで表現したものである。また、GTPと対話行為の情報の有用性を比較するために、発話時間や発話ターン数などの発話情報を用いた推定モデルと単純な視線情報（議論の際に話し手や聞き手を見ている時間）を用いた推定モデルを構築した。さらに、マルチモーダルの有効性を評価するために、GTPと対話行為を用いた推定モデルとGTP、対話行為、発話情報、単純な視線情報を用いた推定モデルを構築した。その結果、多人数での議論における個人の共感スキルレベルの推定には、話者交替/継続時の視線行動と対話行為の情報が有効であることが示唆された。これより、話者交替/継続時の詳細な視線行動を特微量とすることで、共感スキルレベルを推定できる可能性を示した。

Tanら[26]は、OMG-Empathy Datasetにおいて、共感している聞き手が、話し手のストーリーを聞いているときの感情の価数を時間経過とともに予測するLong Short Term Memory (LSTM) モデルを提案している。OMG-Empathy Datasetは、1人が演者（話し手）で、もう1人が参加者（聞き手）である2者間のインタラクションで構成されている。話し手は4名、聞き手は10名である。話し手は架空の自伝的な物語（幼なじみの話や飛行機で嫌な思いをした話など）を聞き手に話し、平均5分12秒の動画を80本作成した。各セッションの後、聞き手は話し手がストーリーを話しているときに自身がどう感じたか（共感的反応）を-1（否定的）～1（肯定的）の範囲で評価した。この研究では、上記のデータセットから音声、言語、視聴覚特微量を抽出し、聞き手の共感反応を予測するLSTMモデル構築している。また、モデルの予測と参加者が自己申告した共感反応との間で一致相関係数（CCC）を用いて評価を行っている。その結果、言語特微量を含むモデルは含まないモデルよりも性能が優れていることが明らかになり、相手の話に共感する際の反応に影響を与える最も重要な情報は相手の話の内容であることが考えられている。

Soleymaniら[27]は、対話中の言語的行動と非言語的行動から自己開示レベルを推定している。人間同士もしくは人間とエージェントの擬似モノログ対話が行われているDistress Analysis Interview Corpus (DAIC)[31]とSocial Anxiety and Self-disclosure Dataset[32]の2つのデータセットを用いている。2つのデータセットは、メンタルヘルスの症状（不安と苦痛）に関する研究で記録されたもので、特定の質問への回答に関する自己開示を外部の評価者がラベル付けしたものである。これらのデータセットから言語、音声、視覚に関する特微量を抽出し、自己開示レベルを推定する回帰タスクを扱っている。その結果、言語特微量に含まれる単語数、言葉で表現される感情的・認知的プロセス、文章の構成が親密な自己開示の重要な指標となることを明らかにしている。加えて、頷きなどのヘッドジェスチャや発話のポーズも自己開示と関連していることを明らかにしている。また、各モダリティのパフォーマンスとロバスト性を分析するために、コーパスごとの評価とコーパス間での評価を行っている。その結果、コーパスごとの評価では、言語特微量が自己開示の推定に最も適していることを明らかにしている。一方で、コーパス間での評価では、非言語（音声、視覚）特微量が安定した性能を発揮することを明らかにしている。言語、非言語を合わせたマルチモーダル特微量は、コーパスごとの評価では最良のモダリティ

(言語)と同等の性能となり, コーパス間での評価では最良のモダリティを上回る性能を示した.

## 第3章 研究課題

本章では、本研究における問題の定義と研究課題について述べる。

### 3.1 問題の定義

褒める行為は、対象の行動や性格に向けられた称賛を表現する言語・非言語行動であると考えられている [1][2]。さらに、褒める行為は、褒める人から褒められる人への一方的な意思の伝達ではなく、複雑な社会的コミュニケーションであり、褒められる人の役割は褒める人の役割と同じくらい重要であると考えられている [3]。

2.1 節で述べたように、教育学、心理学の分野における褒める行為に関連する研究事例の多くは、アンケート調査や特定のタスク後に実施した評価の結果を用いて分析を行い、褒める行為の効果や影響を明らかにしている。Anderson ら [10] は、絵画を描くという課題を用いて実験を行い、ポジティブな言語的報酬（例：その絵は本当にいいね）を与えられた場合は絵を描く時間が長くなり、受賞的な報酬や金銭的な報酬の場合を与えられた場合は絵を描く時間が短くなることを明らかにしている。褒める行為による効果は確認できるが、言語・非言語行動をどのように用いて褒めると効果があるのかは明らかにされていない。Stipek ら [11] は、パズルを成功させる課題に取り組んでもらったあと、「本当に上手にパズルができたね」のように成功したことについて褒めるもしくは「パズルが終わったね」のように成功した事実についてコメントするという実験を行い、成功したことについて褒められた参加者は成功した事実についてコメントされた参加者よりも笑顔や堂々としたポーズなどのポジティブな表出を多くすることを明らかにしている。褒める行為における効果は確認できるが、褒める人は言語・非言語行動をどのように用いて褒めると効果があるのか、褒められる人がどの程度の非言語行動が確認されたのか、工学的な分析は行われていない。

加えて、2.2 節で述べたように、特定のタスクやシーンにおける言語・非言語行動から、個人の性格特性やコミュニケーションスキル、共感スキルといった能力やパフォーマンスを推定する取り組みが行われている。Okada ら [22] は、人事管理の経験を持つ外部評価者によって評価された個人のコミュニケーションスキルを推定する機械学習モデルを構築し、コミュニケーションスキルを最もよく推定することができる言語・非言語行動の組み合わせ分析している。Tan ら [26] は、共感している聞き手が話し手のストーリーを聞いているときの感情の価数を予測する LSTM モデル構築し、言語特徴量を含むモデルは言語特徴量を含まないモデルよりも性能が優れていることを明らかにしている。コミュニケーションスキルや共感スキルは褒める行為に関連のあるスキルであると考えられるが、褒める行為に関する分析は行われていない。また、本研究が属する研究領域において、特定のタスクやシーン中の褒める際の言語・非言語行動を分析する研究は行われておらず、言語・非言語行動をどのように用いると上手く褒めることができるのか明らかにされていない。

上記をふまえ、本研究における問題は、言語・非言語行動をどのように用いると上手く褒めることができるのか明らかにされていないことであると定義できる。

## 3.2 研究課題の設定

3.1節で述べたように、言語・非言語行動をどのように用いると上手く褒めることができるのか明らかにされていないという問題があり、相手を上手く褒めることが苦手な人は何を意識すれば褒め方が上達するのか分からないということが考えられる。この問題を解決するためには、相手を上手く褒めるために言語・非言語行動をどのように用いると良いか明らかにする必要がある。そこで、人間の言語・非言語行動を利用して、対話における褒め方の上手さを分析する取り組みを行う。これにより、対話において相手を上手く褒めるためにはどのような行動が重要であるか明らかになり、相手を上手く褒めるための能力の向上や褒め方の上手さを評価するシステムの開発が期待される。

本研究では、対話における褒め方の上手さと人間の行動の関係を分析する取り組みとして、話者（褒める人・褒められる人）の頭部と顔部の振舞いに着目し、相手を上手く褒めるためにはどのような行動が重要であるのか明らかにする取り組みを行う。既存知見より、褒める行為は褒める人から褒められる人への一方的な意思の伝達ではなく、褒められる人の役割は褒める人の役割と同じくらい重要であると考えられていることから [3]、褒める人だけでなく褒められる人にも着目する。また、頭部と顔部の振舞いは意図や感情、反応を表現することができ、コミュニケーションにおいて重要であると考えられていることから [33]、言語・非言語行動の中から非言語行動である頭部と顔部の振舞いに着目する。

さらに、本研究では、先行研究 [4][5][6][7] の改善を行う。1つ目は、対話データの拡張を行う。先行研究 [4][5][6] では、20名分の対話データ（計150分）を対象に、褒め方の上手さの評価値を推定する取り組みを行ってきた。しかし、上手く褒めるための頭部・顔部の振舞いを分析するためには、対話データ数が少ないことが考えられる。そこで本研究では、14名分の対話データ（計105分）を追加し、合わせて34名分の対話データ（計255分）を対象に、褒め方の上手さの評価値を推定する取り組みを行う。2つ目は、発話シーンの定義を改良する。先行研究 [4][5][6] では、発話シーンを200ミリ秒未満の連続した音声区間と定義をしていた。しかし、発話が細かく区切れることが確認できた。そこで本研究では、自然な発話の区切れとなるように、対話データを確認した上で、発話シーンを400ミリ秒未満の連続した音声区間と定義する。3つ目は、アノテータの人数とアノテーション方法を改良する。先行研究 [4][5][6] では、対話に参加していない1名のアノテータが褒めていると判断できる発話シーンを判定し、対話において褒められる人が対話相手（褒める人）の褒め方の上手さを評価していた。しかし、褒めているシーンの判定や褒め方の上手さを評価を1名で行っていることから、主観的な結果となる可能性が考えられた。複数名のアノテータによる人間のスキルやパフォーマンスを評価する研究事例は多く存在する。そこで本研究では、対話に参加していない5名のアノテータが褒めていると判断できる発話シーンの判定と褒める人の褒め方の上手さの評価を行うようにする。加えて、アノテータが褒めていると判断できる発話シーンを Praise シーン、褒められる人が評価した褒め方の上手さの結果を Praise スコアと定義していた。本研究では、アノテーション方法を改良したことに伴い、5名のアノテータのうち3名以上のアノテータが褒めていると判断できる発話シーンを Praise シーン、褒めていると判断したアノテータの評価の平均

値を Praise スコアと定義する。4つ目は、機械学習手法を変更する。先行研究 [4][5][6] では、褒め方の上手さの評価値を Gradient boosting を用いて回帰モデルを構築し、褒め方の上手さの評価値を推定していた。しかし、評価値の分布に偏りがあることから、推定精度（決定係数）が  $R^2 = 0.236$  にとどまっていた。評価値を低、中、高などのクラスに分け、分類モデルを構築し評価値が属するクラスを推定している研究事例もある。そこで本研究では、研究の初期段階として、Random Forests を用いて分類モデルを構築し、褒め方の上手さの評価値が属するクラスを推定する。5つ目は、上手く褒めるための振舞いを明らかにする。先行研究 [7] では、機械学習モデルの結果から上手く褒めるために重要である振舞いを明らかにしてきた。しかし、褒め方の上手さの評価値とどのような関連があるのか明らかにできていなかった。そのため、上手く褒めるためには振舞うとよいのか、上手く褒めるためにはそこで本研究では、上手く褒めるために重要と判断した振舞いが、褒め方の上手さの評価値とどのような関連があるのか明らかにする。

上記をふまえ、本研究では、褒める人と褒められる人の頭部、顔部の振舞いに着目し、相手を上手く褒めるためにはどのような振舞いが重要であるのか明らかにすることを研究課題として設定する。

## 第4章 対話コーパス

本章では、本研究で用いる対話コーパスについて述べる。本研究では、2者対話における話者の頭部・顔部の振舞いと、褒め方の上手さに関する評価値を含む対話コーパスの作成を行った [4][5][6][7]。具体的には、話者の頭部・顔部の振舞いを記録するために2者対話の収録を行い、その後対話データのアノテーションと褒め方の上手さに関する評価を行った。

## 4.1 2者対話の収録

本節では、2者対話の収録について述べる。頭部・顔部の振舞いを記録するために2者対話の収録を行った。2者対話の参加者は、20代の大学生34名（男性28名，女性6名）であり、2名1組のペアを17組構成した。17組のうち、初対面が14組，顔見知り2組，友人同士が1組であった。対話の収録を始めるにあたり，参加者に対話材料を準備させることを意図して，いままで頑張ってきたことに関するエピソードを2つ以上用意してもらった。対話収録時は，図4.1のように参加者が互いに向き合って着座してもらった。このと



図 4.1: 2者対話の様子

き参加者間の距離は180cmとした。対話の記録は，各参加者の様子と2者対話全体の様子を撮影するためのビデオカメラ，各参加者の声を録音するためのマイクを用いて行った。各組の参加者（参加者A，参加者B）は，撮影者の合図に従い，次の(1)～(3)を行った。

- (1) 自己紹介（5分間）
- (2) 参加者Aが褒める人となり，参加者Bが褒められる人となる対話（5分間）

(3) 参加者 B が褒める人となり，参加者 A が褒められる人となる対話（5 分間）

(1)~(3) の対話を 17 組分，計 255 分間収録した．なお，(1) の自己紹介は，各組の多くが初対面であることから，参加者の緊張をほぐす目的で行っているため，分析の対象外とする．(2) と (3) の対話において，褒める人には，対話相手を積極的に褒めるように指示した．しかし，一方的に褒めているだけのような不自然な対話にならないようにするために，自由に質問したり，リアクションをしたりすることを許可した．褒められる人には，事前に用意した自分がいままで頑張ってきたことに関連するエピソードを話すように指示した．加えて，対話の自然さや話題の多様性を担保するために，事前に用意していないエピソードについて話すことを許可した．

## 4.2 アノテーション

本節では，アノテーションと褒め方の上手さの評価について述べる．4.1 節で収録した対話データのアノテーションと褒め方の上手さに関する評価を行った．はじめに，映像データや音声データに対して注釈付けを行うツールである ELAN[34] を用いて，各参加者の映像データと音声データに対して発話シーンを人手で付与した．発話シーンの件数と時間を表 4.1 に示す．発話シーンは，無音区間が 400 ミリ秒未満の連続した音声区間であ

表 4.1: 発話シーンと Praise シーンの情報

	件数	平均値	最大値	最小値
発話シーン（褒める人）	2701 件	1.324 秒	23.117 秒	0.062 秒
発話シーン（褒められる人）	3413 件	2.040 秒	26.234 秒	0.018 秒
Praise シーン	228 件	2.018 秒	9.127 秒	0.368 秒

る．先行研究 [4][5][6] では無音区間が 200 ミリ秒未満の連続した音声区間\*としていたが，発話が細かく区切れることが確認できたため，本稿では発話の区切りが自然になるように無音区間が 400 ミリ秒未満の連続した音声区間とした [7]．

次に，人間が能力やパフォーマンスを評価することは広く行われていることから [35]，2 者対話の収録に参加していない第三者のアノテータ 5 名が，褒め方の上手さの評価を行った．本研究では，褒める人の正面に設置したビデオカメラから撮影した映像データと，褒める人に取り付けたマイクから録音した音声データを参照し，褒める人の発話シーンごとに下記の判定・評価を行った．

- 対話の相手を褒めているシーンであるか，そうでないかの判定

\*国立国語研究所「日本語話し言葉コーパス」において，転記基本単位（言語音が 0.2 秒以上の途切れがなく，連続して生じている区間）とされている

- 褒めているシーンであると判断した場合，1（上手く褒められていない）～7（上手く褒められている）の7段階での褒め方の上手さの評価

本稿では，各発話シーンにおいて，褒めているシーンであると判定したアノテータが3名以上であるシーンを Praise シーンとする．Praise シーンの件数と時間を表 4.1 に示す．また，各 Praise シーンにおいて，褒めているシーンであると判定したアノテータの評価の平均値を Praise スコアとする．Praise スコアの分布を図 4.2 に示す．アノテータ間の Praise

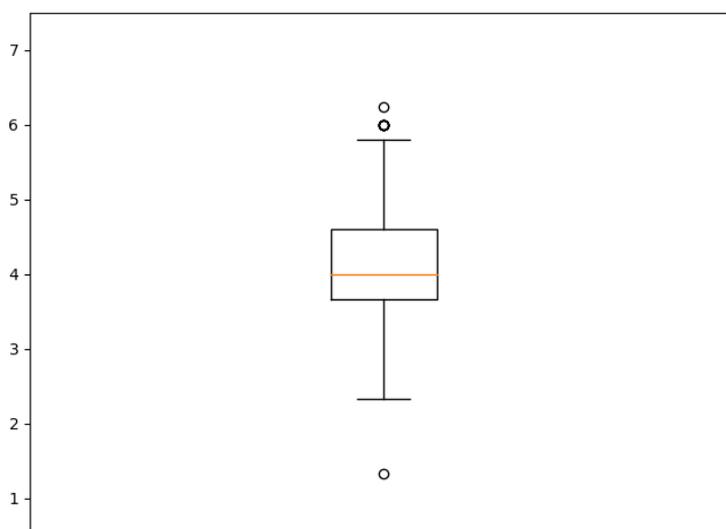


図 4.2: Praise スコアの分布 (N=228)

スコアの一貫性を評価するために，級内相関係数 (ICC) [36] を利用した．Praise スコアの一貫率は，3～5 人のアノテータのそれぞれの組合せごとに級内相関係数を算出した後，サンプル数を考慮して重み付き平均を算出した結果， $ICC(2, k) = 0.571$  であった．この結果から，Praise スコアはアノテータ間で中程度の一貫率があり，信頼性のあるデータであることが示唆された．

先行研究 [4][5][6] では，対話に参加していない 1 名のアノテータが発話シーンと Praise シーンを判定し，対話において褒められる人が対話相手（褒める人）の褒め方の上手さを評価していた．しかし，Praise シーンの判定や褒め方の上手さの評価を 1 名で行っていることから，主観的な結果となる可能性が考えられた．そこで，対話に参加していない 5 名のアノテータが褒めていると判断できる発話シーンの判定と褒める人の褒め方の上手さを評価を行うようにした．そのため，褒める行為を直接受けている本人と，客観的に確認している第三者で，印象が異なるのか確認する必要がある．そこで，褒め方の上手さについて，褒められた本人とアノテータの評価に違いがあるか確認するために，対話参加者

に褒め方の上手さの評価を行った。具体的には、対話参加者5名に自身が褒められる人であった対話における褒める人の映像データと音声データを参照してもらい、褒める人の発話シーンごとに下記の判定・評価を行った。

- 対話の相手を褒めているシーンであるか、そうでないかの判定
- 褒めているシーンであると判断した場合、1（上手く褒められていない）～7（上手く褒められている）の7段階での褒め方の上手さの評価

対話参加者5名が判定・評価した結果と2者対話の収録に参加していない第三者のアノテータ5名が判定・評価した結果の相関係数を確認した。その結果、相関係数が0.773であり、対話参加者が評価した場合でも、アノテータが評価した場合でも、褒める人の褒め方に対する印象におおむね違いがないことが確認できた。

## 第5章 上手く褒めるための頭部・顔部の 振舞いの分析

本章では、上手く褒めるための頭部・顔部の振舞いの分析について述べる。本研究では、対話中の褒める人と褒められる人のどのような非言語行動が相手を上手く褒めるために重要であるのか明らかにする取り組みを行う。そのため、褒め方の上手さを推定する機械学習モデルを構築し、どのような行動が重要であるのか判断する必要がある。具体的には、頭部、視線、表情に関連する特徴量を抽出し、褒め方の上手さの評価値を推定する機械学習モデルを構築する。さらに、機械学習モデルにおいて重要であると判断された特徴量に関連する振舞いについて分析を行う。

## 5.1 特徴量抽出

本節では、頭部、視線、表情に関する特徴量の抽出について述べる。マルチモーダルインタラクションの分野では、頭部、視線、表情といった非言語行動に関する特徴量が行動分析に重要な役割を果たすことが多くの研究で示されている。さらに、非言語行動は意図や感情、反応を表現することができ、コミュニケーションにおいて重要であると考えられている [33]。そこで本研究では、非言語行動の中から頭部、視線、表情に着目する。頭部は、例示的動作、調整的動作、情動表出、適応動作に分類できるとされており、コミュニケーションにおいて言語行動の補完や相手への合図（肯定/否定の動作や傾聴の動作）の機能を果たすと考えられている [37]。視線は、身体行動には表れない感情、態度、関係に関する信号を提供すると考えられており、コミュニケーションにおいて多くの機能を果たすと考えられている [38]。表情は、人間の感情と密接に関係しており、自分自身、相手、他者に対する感情に関する豊富な情報を提供する手がかりであると考えられている [39]。

本研究では、顔画像処理ツールである OpenFace[40] を用いて、対話の参加者（褒める人と褒められる人）の正面に設置したビデオカメラで撮影した映像データから頭部、視線、表情（Action Units[28]）に関する特徴量を抽出した。抽出した特徴量は次のとおりである。

**頭部** ビデオカメラ側から顔を見て、左から右方向を x 軸、下から上方向を y 軸、手前から奥方向を z 軸とした場合（図 5.1）、Praise シーンの前後 1 秒ずつを含む範囲（図 5.2）における頭部の x 軸、y 軸、z 軸周りの回転角度（pose\_Rx, pose\_Ry, pose\_Rz）の分散（\_var）、中央値（\_med）、10 パーセンタイル値（\_p10）、90 パーセンタイル値（\_p90）を用いた。

**視線** ビデオカメラ側から顔を見て、左から右方向を x 軸、下から上方向を y 軸とした場合、Praise シーンの前後 1 秒ずつを含む範囲における視線の x 軸、y 軸方向の角度（gaze\_Ax, gaze\_Ay）の分散、中央値、10 パーセンタイル値、90 パーセンタイル値を用いた。

**Action Units** Action Units は、人間の顔部における筋肉群の基本的な行動の単位を表している [28]。Praise シーンの前後 1 秒ずつを含む範囲における OpenFace で用い

られている各 Action Units (表 5.1) の強度の分散, 中央値, 10 パーセンタイル値, 90 パーセンタイル値を用いた。

これらの特徴量は, 対話において褒める人 (praiser\_) と褒められる人 (receiver\_) からそれぞれ抽出した。また, 特徴量の取りうる値の範囲を揃えるために, 各特徴量を平均値 0, 分散 1 となるように正規化を行った。

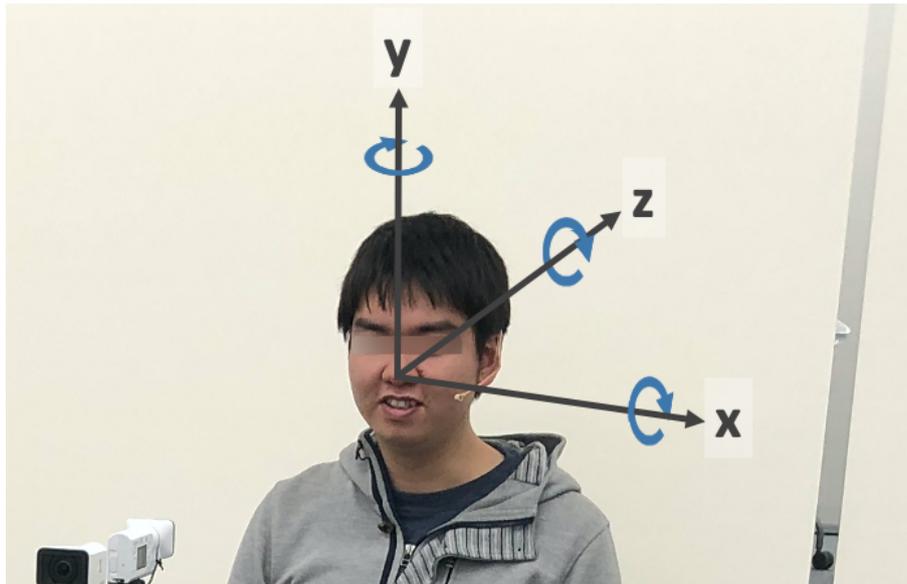


図 5.1: 頭部の x 軸, y 軸, z 軸方向の座標系

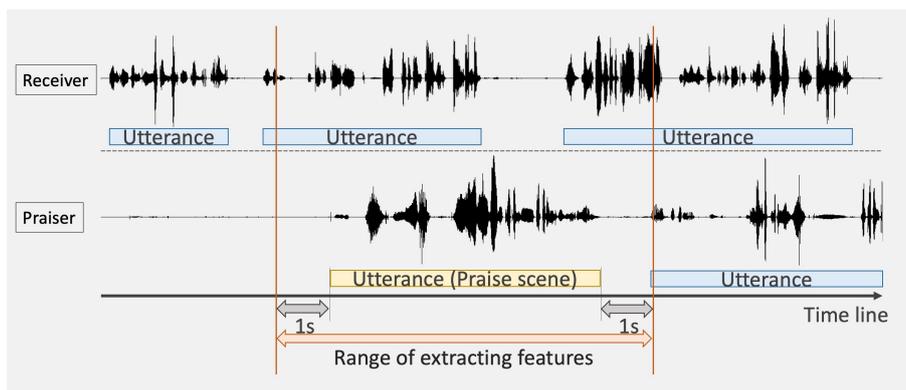


図 5.2: 特徴量を抽出する範囲

表 5.1: Action Units の内容

項目	内容	項目	内容
AU01	眉の内側を上げる	AU14	笑窪を作る
AU02	眉の外側を上げる	AU15	唇の両端を下げる
AU04	眉を下げる	AU17	顎を上げる
AU05	上瞼を上げる	AU20	唇の両端を横に引く
AU06	頬を持ち上げる	AU23	唇を固く閉じる
AU07	瞼を緊張させる	AU25	顎を下げずに唇を開く
AU09	鼻に皺を寄せる	AU26	顎を下げて唇を開く
AU10	上唇を上げる	AU45	瞬きをする
AU12	唇の両端を引き上げる		

## 5.2 褒め方の上手さの推定モデルの構築

本節では、4.2節で評価した褒め方の上手さの評価値（Praise スコア）と、5.1節で抽出した頭部、視線、表情に関する特徴量を用いて、褒め方の上手さを推定する機械学習モデルについて述べる。本研究では、頭部と顔部の振舞いに関連する特徴量を利用して、どのような振舞いを行うと上手く褒められるのか判断するために、褒め方の上手さを推定する機械学習モデルを構築した。具体的には、Praise シーンを Praise スコア低群、中群、高群の3クラスに分け、頭部と顔部の振舞いに関する特徴量に基づいて当該シーンにおける Praise スコアがどのクラス（低、中、高群）に属するか推定する分類器を作成した。

各群に属する Praise シーン数ができるだけ均等になるように、下記のように Praise スコア低～高群を定義した\*。

- Praise スコア低群：Praise スコアが3.8点以下の Praise シーン（計82シーン）
- Praise スコア中群：Praise スコアが3.8点より大きく、4.4点未満の Praise シーン（計65シーン）
- Praise スコア高群：Praise スコアが4.4点以上の Praise シーン（計81シーン）

推定モデルを構築する手段として、特徴量の重要度を評価することができる Random forests[41]を用いた。木の本数や木の深さなどのハイパパラメータは、Hyperopt[42]でチューニングした。データセットを90%の訓練データ、10%のテストデータに無作為に分け、訓練データで構築したモデルを用いてテストデータが属するクラスを推定するタスクを100回行った。

\*各群に属するシーン数は等しいことが理想ではあるが、スコアが同じシーンが多数存在したため、各群のシーン数を等しくすることができなかった。

表 5.2: 各指標の平均値 (N=100)

	適合率	再現率	F 値
Baseline	0.367	0.340	0.341
Model P	0.459	0.446	<b>0.433</b> *†
Model R	0.452	0.405	0.391*
Model P+R	0.509	0.492	<b>0.467</b> *‡

### 5.3 褒め方の上手さの推定モデルの結果

構築した各モデルの指標の平均値を表 5.2 に示す。Baseline (チャンスレベル) は、データセットにおける各群の割合にあわせて、36%、28%、36%の確率で Praise スコアの低、中、高群を出力するモデルを用いた。本研究で提案するモデルは、Praise スコアを低、中、高の3群を予測するモデルであり、褒める人 (Praiser) から抽出した特徴量を用いたモデル (Model P)、褒められる人 (Receiver) から抽出した特徴量を用いたモデル (Model R)、褒める人と褒められる人から抽出した特徴量を用いたモデル (Model P+R) を用いた。各モデル間での差を確認するために Baseline と提案するモデルとの間において Bonferroni 補正を用いて対応のある t 検定を行った。その結果、Baseline と提案するモデル間で 1% 水準で有意差が認められることが確認できた。このことから、本研究で提案するモデル (Model P, Model R, Model P+R) は、Baseline よりも性能を向上させることができたと判断できる。さらに、提案するモデル間において Bonferroni 補正を用いて対応のある t 検定を行った。その結果、Model P と Model R 間において 5% 水準で有意差が認められ、Model R と Model P+R 間において 1% 水準で有意差が認められることが確認できた。このことから、褒める人に関する情報と褒められる人の情報を用いることにより、褒め方の上手さの評価値を推定するモデルの性能を向上させることができ、ある程度妥当に Praise スコアと特徴量の関係をモデル化できたとと言える。

### 5.4 上手く褒めるための頭部・顔部の振舞いの分析

本節では、上手く褒めるための頭部・顔部の振舞いの分析について述べる。具体的には、5.2 節で構築した機械学習モデルにおいて、重要度が高いと判断された特徴量について分析を行う。表 5.2 において推定精度が高かった Model P における特徴量の重要度上位 10 件を図 5.3、Model P+R における特徴量の重要度上位 10 件を図 5.4 に示す。本研究では、Gain<sup>†</sup>を用いて特徴量の重要度を算出した。図 5.3 と図 5.4 において、橙色は褒める人に関する特徴量、青色は褒められる人に関する特徴量を表している。

\*Baseline との間で  $p < .01$ †Model R との間で  $p < .05$ ‡Model R との間で  $p < .01$ 

†特徴量の分岐により得た目的変数の改善幅の評価

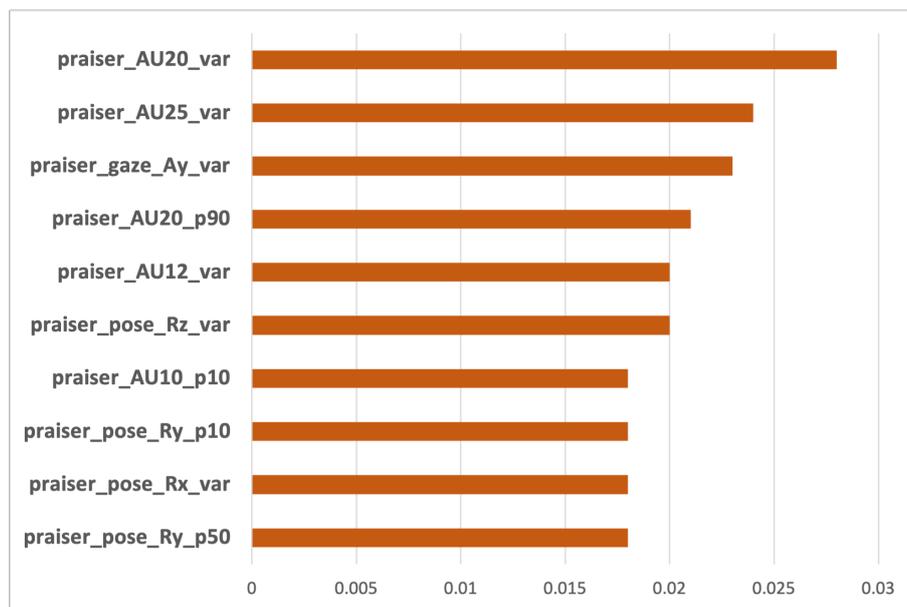


図 5.3: Model P における特徴量の重要度上位 10 件

#### 5.4.1 褒める人の頭部・顔部の振舞いの分析

本項では、褒める人の頭部・顔部の振舞いの分析を行い、上手く褒めるためにはどのような振舞いが重要であるのか明らかにする。褒める人について、Model P (図 5.3) と Model P+R (図 5.4) において gaze\_Ay, pose\_Rz, AU12, AU20, AU25 に関連する特徴量の重要度が高かった。それぞれ、gaze\_Ay は視線の y 軸方向 (上下方向) の回転角度、pose\_Rz は頭部の z 軸回り (奥行き方向) の回転角度、AU12 は唇の両端を引き上げる動き、AU20 は唇の両端を横に引く動き、AU25 は顎を下げずに唇を開く動きを表している。図 5.3 と図 5.4 より、gaze\_Ay, pose\_Rz, AU12, AU20, AU25 の分散を表す特徴量が、重要度の高い特徴量であることがわかる。これより、褒める人は、視線を上下に動かす、頭部を前後に動かす、唇の両端を引き上げる、唇の両端を横に引く、顎を下げずに唇を開く振舞いが、褒めている最中に動かしていることが考えられる。加えて、図 5.3 と図 5.4 より、AU20 の 90 パーセンタイル値を表す特徴量が、重要度の高い特徴量であることがわかる。これより、褒める人は、唇の両端を横に引く振舞いが褒めている最中に強く表れていることが考えられる。上記より、上手く褒めるためには視線を上下に動かす、頭部を前後に動かす、唇の両端を引き上げる、唇の両端を横に引く、顎を下げずに唇を開く振舞いの変化や唇の両端を横に引く振舞いの強度に注目するとよいことが考えられる。

次に、上手く褒めるため注目するとよいと判断した振舞いの中で、上手く褒めるためにどの振舞いが重要であるのか明らかにする。具体的には、重要度の高かった特徴量について、Praise スコア高群と Praise スコア低群の間で振舞いに違いがあるのか分析を行う。Praise スコア高群・低群における重要度の高かった特徴量の中央値を表 5.3 に示す。これら

\*\*Praise スコア高群と Praise スコア低群との間で  $p < .01$

\*Praise スコア高群と Praise スコア低群との間で  $p < .05$

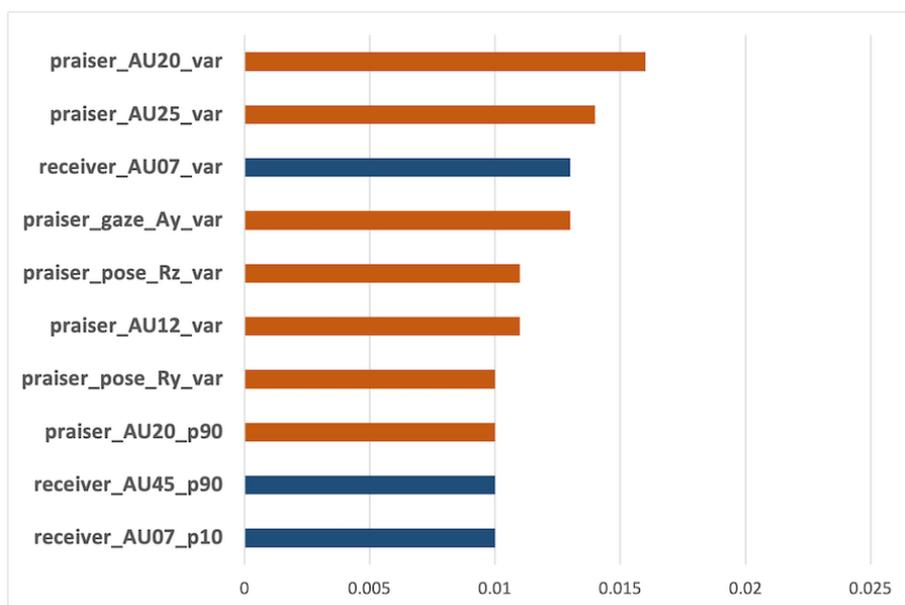


図 5.4: Model P+R における特徴量の重要度上位 10 件

の特徴量について、Praise スコア高群と Praise スコア低群の間で対応のない t 検定を行ったところ、praiser\_gaze\_Ay\_var (gaze\_Ay の分散値)、praiser\_AU12\_var (praiser\_AU12 の分散値)、praiser\_AU25\_var (praiser\_AU25 の分散値) が Praise スコア低群よりも Praise スコア高群で有意に大きいことが確認できた。これはすなわち、Praise スコア高群において、視線を上下に動かす、唇の両端の引き上げる、顎を下げずに唇を開く振舞いが褒めている最中により多く変化していることが考えられる。実際の対話コーパス中でこれらの動作が行われているシーンを図 5.5 に示す。これより、褒めている最中に、視線を上下に動かす、唇の両端の引き上げる、顎を下げずに唇を開く振舞いが変化していることが確認できる。相手を上手く褒めるためには、視線を上下に動かす、唇の両端の引き上げる、顎を

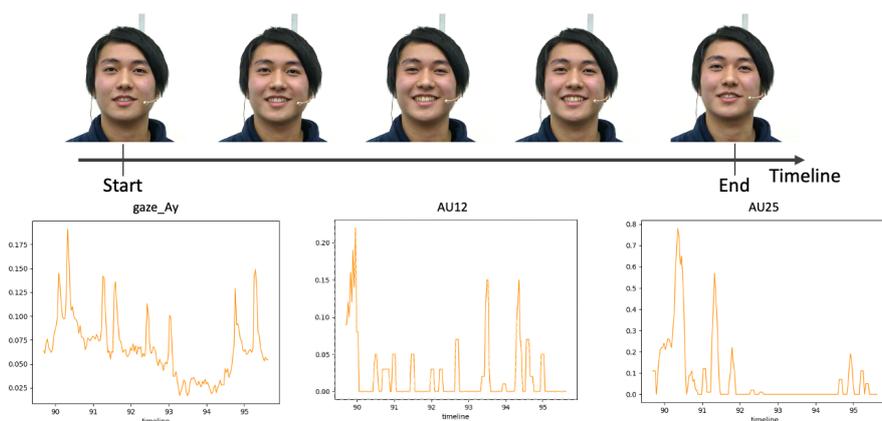


図 5.5: 実際の対話シーン (Praise シーン) の画像と特徴量の値の一例

表 5.3: Praise スコア高群と低群における重要度の高い特徴量の中央値

特徴量	Praise スコア高群	Praise スコア低群
praiser_AU20_var	0.025	0.006
praiser_AU25_var**	<b>0.436</b>	0.248
receiver_AU07_var	0.161	0.154
praiser_gaze_Ay_var*	<b>0.002</b>	0.001
praiser_pose_Rz_var	0.001	0.000
praiser_AU12_var*	<b>0.204</b>	0.082
praiser_pose_Ry_var	0.001	0.000
praiser_AU20_p90	0.298	0.277
receiver_AU45_p90	1.020	1.008
receiver_AU07_p10	1.532	1.532

下げずに唇を開く振舞いをより動かすことが重要であると考えられる。

最後に、上手く褒めるために重要であると判断した行動が、Praise スコアとどの程度の関連性があるのか確認するために、重要度が高い特徴量と Praise スコアの間のピアソンの相関係数を算出した。その結果を表 5.4 に示す。表 5.4 より、praiser\_AU25\_var (praiser\_AU12 の分散値) において弱い正の相関が確認できた。これより、上手く褒めるために重要であると判断した行動の中で、顎を下げずに唇を開く振舞いは Praise スコアが高くなるにつれて、振舞いの変化が大きくなることが考えられる。図 5.5 より、褒め始めと褒め終わり、顎を下げずに唇を開く振舞いをしていることが確認できる。上手く褒めるためには、顎を下げずに唇を開く振舞いを行うことが最も重要であると考えられる。

#### 5.4.2 褒められる人の頭部・顔部の振舞いの分析

本項では、褒められる人の頭部・顔部の振舞いの分析を行い、褒め方の上手さにどの程度の関連性があるのか明らかにする。褒める行為は、褒める人から褒められる人への一方的な意思の伝達ではなく、複雑な社会的コミュニケーションであり、褒められる人の役割は褒める人の役割と同じくらい重要であると考えられていることから [3]、褒められる人の頭部・顔部の振舞いと褒め方の上手さがどの程度の関連性があるのか確認する。褒められる人について、Model P+R (図 5.4) において AU07, AU45 に関連する特徴量の重要度が高かった。AU07 は瞼を緊張させる動き、AU45 は瞬きをする動きを表している。図 5.4 より、AU07 の分散、AU07 の 10 パーセンタイル値、AU45 の 90 パーセンタイル値を表す特徴量が、重要度の高い特徴量であることがわかる。これより、褒められる人の瞼を緊張させる動きの変化や強度、瞬きをする動きの強度が褒め方の上手さと関連があるのではないかと考えられる。次に、褒め方の上手さと関連があると判断した振舞いが、褒め方の上手さとどの程度の関連性であるのか確認する。そのために、Praise スコア高群と

表 5.4: 重要度が高い特徴量と Praise スコアの相関係数

特徴量	相関係数	p 値
praiser_AU20_var	0.063	0.346
praiser_AU25_var	0.216	0.001
receiver_AU07_var	-0.009	0.892
praiser_gaze_Ay_var	0.108	0.104
praiser_pose_Rz_var	0.037	0.579
praiser_AU12_var	0.161	0.015
praiser_pose_Ry_var	0.065	0.330
praiser_AU20_p90	0.100	0.131
receiver_AU45_p90	0.008	0.909
receiver_AU07_p10	-0.091	0.173

Praise スコア低群の間で対応のない t 検定を行った。その結果、表 5.3 より、Praise スコア高群と Praise スコア低群の間で有意な差は確認できなかった。さらに、重要度が高い特徴量と Praise スコアの間でピアソンの相関係数を算出した。その結果、表 5.4 より、重要度が高い特徴量と Praise スコアの間に相関は確認できなかった。上記より、褒められる人の顔を緊張させる動きの変化や強度、瞬きをする動きの強度が褒め方の上手さと関連があるのではないかと考えられたものの、これらの振舞いが褒め方の上手さとどの程度の関連性があるかまでは確認できなかった。本研究では、褒める人と褒められる人に関する特徴量を同時に抽出している。すなわち、褒められる人に関する振舞いは、相手に褒められている最中の振舞いを扱っている。しかし、褒める人の褒め方の上手さに関連する褒められる人の振舞いは、褒められた後に表れることも考えられる。既存知見より、褒められた参加者は、笑顔や堂々としたポーズなどのポジティブな表出を多くすることが明らかにされている [11]。今後、振舞いに関連する特徴量の抽出する範囲のパターンを増やすことで、褒められる人のどのような振舞いが褒める人の褒め方の上手さに関連しているのか分析を行う必要がある。

## 5.5 本分析で得られた知見

5.4 節の分析で得られた知見は次のとおりである。

- 上手く褒めるためには視線を上下に動かす、頭部を左右に動かす、頭部を前後に動かす、唇の両端を引き上げる、唇の両端を横に引く、顎を下げずに唇を開く振舞いの変化や唇の両端を横に引く振舞いの強度に注目するとよいことが考えられること。
- 相手を上手く褒めるためには、視線を上下に動かす、唇の両端の引き上げる、顎を下げずに唇を開く振舞いをより動かすことが重要であると考えられ、その中でも顎

を下げずに唇を開く振舞いを行うことが最も重要であると考えられること。

本研究では上記に示す知見が得られたが、分析を行う上でいくつかの制約がある。1つ目に、本研究では、褒める人が実際に褒めているシーンの振舞いを分析しており、褒める人が褒めていないシーン（例：発話シーンではあるが、褒めていると判断されていないシーン）における振舞いは分析できていない。そのため、褒め方の上手さを判断するために着目する振舞いが、褒めていないシーンではどのように扱われているのかは確認する必要がある。今後、褒めていないシーンでの振舞いも分析し、その分析をふまえて褒めるときはどのような振舞いが表出するのか明らかにする取り組みを行う必要がある。

2つ目に、本研究では研究の初期段階として、先行研究で使用した Gradient boosting[43]ではなく、Random Forests を用いて分類モデルを構築し、褒め方の上手さの評価値が属するクラスを推定した。今後、先行研究で使用した Gradient boosting を用いて褒め方の上手さが推定できるのか分析する必要がある。さらに、SVMなどの他の機械学習手法を含め、どの機械学習手法が褒め方の上手さを推定することに適しているのか検証をする必要がある。

3つ目に、本研究では、母国語が日本語の参加者を対象に言語・非言語行動を記録した。そのため、本研究で明らかにした知見は日本人もしくは日本の文化に精通している人には有用であると考えられる。しかし、言語・非言語行動は文化、宗教、価値観の違いによって、同じ意味でも異なる行動を示すことが考えられる [44]。例えば、日本では頭を縦に振り肯定 (Yes) を表現するが、ブルガリアでは頭を横に振り肯定 (Yes) を表現している。また、日本では頭を横に振り否定 (No) を表現するが、ギリシャでは頭を縦に振り否定 (No) を表現している。このように、ひとつの振舞いでも、文化によって異なる意味を表している。今後、本研究で明らかにした知見が日本とは異なる文化、宗教、価値観を持つ人々に有用であるか検証する必要がある。

## 第6章 結論

本研究では、対話における褒め方の上手さと人間の行動の関係を分析する取り組みとして、話者(褒める人・褒められる人)の頭部と顔部の振る舞いに着目し、相手を上手く褒めるためにはどのような行動が重要であるのか明らかにする取り組みを行った。はじめに、頭部と顔部の振舞いと、褒め方の上手さの評価値を含む対話コーパスを作成した。次に、頭部と顔部の振舞いに関連する特徴量を用いて褒め方の上手さの評価値を推定する機械学習モデルを構築し、どのような頭部と顔部の振る舞いが重要であるかを検証した。さらに、本研究では先行研究[6]の改善を行った。まず、対話データ数を14名分増やし、計34名分の対話データを対象に褒め方の上手さの分析を行った。次に、複数の評価者による褒め方の上手さの評価を行い、その評価結果を用いて褒め方の上手さを推定するタスクを行った。上記の取り組みの結果、褒める人の視線の方向に関する振る舞い、口角や口の開きに関する振る舞いが相手を上手く褒めるためには重要であることが明らかになった。また、褒められる人の顔を緊張させる動きの変化や強度、瞬きをする動きの強度が褒め方の上手さと関連があるのではないかと考えられたものの、これらの振舞いが褒め方の上手さとどの程度の関連性があるかまでは確認できなかった。今後、振舞いに関連する特徴量の抽出する範囲のパターンを増やしたり、褒めていないシーンと比較したりして、褒められる人のどのような振舞いが褒める人の褒め方の上手さに関連しているのか分析を行う必要がある。

本研究にはいくつかの制約がある。1つ目に、本研究では人間の言語・非言語行動の中から、頭部、視線、表情に関する振る舞いに着目した。そのため、言語、音声、ジェスチャ等の言語・非言語行動が上手く褒めるために重要であるのか明らかにできていない。現在、言語、音声、ジェスチャ等を用いた分析の準備を進めている。今後、本研究で取り扱った頭部、視線、表情に関する振る舞いと合わせて、上手く褒めるために重要な行動を明らかにする取り組みを行いたい。2つ目に、本研究では、先行研究から対話コーパスを拡張やアノテーション方法を改良し、20代の大学生34名分の対話データ(計255分間)を対象として褒め方の上手さの分析を行った。しかし、対話コーパスを拡張することはできたが、それでも対話データ数が少ないことが考えられるため、対話コーパスを拡張やアノテーション方法を改良したことの効果を明らかにすることはできなかった。今後、本研究で実施した対話コーパスの拡張やアノテーション方法の改良の効果を明らかにする必要がある。加えて、対話参加者を20代の大学生と限定しているため、本研究で明らかにした知見が年代、性別に関係なく重要であるか検証する必要がある。今後、対話参加者の年代を20代~50代まで均等に集めた上で対話データ数を増やし、本研究で明らかにした知見が年代、性別に関係なく重要であるか検証を行いたい。また、今回の対話コーパスはできるだけ一般的なシーンを想定しているが、個別シーン(例:ビジネス, 教育)においてもこの知見が成立するかは明らかではない。参加者同士の関係性(例:上司と部下, 教師と生徒)によっても、結果は変化する可能性がある。今後、各シーン・関係性に応じて、上手い褒め方に必要な要素にどのような違いがあるか明らかにする必要がある。3つ目に、本研究では、母国語が日本語の参加者を対象に言語・非言語行動を記録した。そのため、本研究で明らかにした知見は日本人もしくは日本の文化に精通している人には有用であると考えられる。しかし、言語・非言語行動は文化、宗教、価値観の違いによっ

て、同じ意味でも異なる行動を示すことが考えられる [44]. 例えば、日本では頭を縦に振り肯定 (Yes) を表現するが、ブルガリアでは頭を横に振り肯定 (Yes) を表現している。また、日本では頭を横に振り否定 (No) を表現するが、ギリシャでは頭を縦に振り否定 (No) を表現している。このように、ひとつの振る舞いでも、文化によって異なる意味を表している。今後、本研究で明らかにした知見が日本とは異なる文化、宗教、価値観を持つ人々に有用であるか検証する必要がある。最後に、本研究では対面での対話を想定しており、本研究で明らかにした結果が遠隔での対話でも同じ結果になるか明らかにできていない。COVID-19の影響で遠隔会議システムを利用したコミュニケーションを行う機会が増えてきている。対面対話と遠隔対話では言語・非言語行動の伝わり方が異なることや、伝達できる情報に変化することが知られている [45]. 特に、遠隔対話では非言語行動が伝わりにくいと考えられる。褒める行為においても、対面対話と遠隔対話では言語・非言語行動の伝わり方や相手を上手く褒めるために重要な行動が異なることが考えられる。これより、遠隔対話において相手を上手く褒めるために重要な行動を明らかにする必要がある。現在、遠隔対話での上手く褒めるために重要な行動を明らかにするための準備を行っている。今後、対面対話のみならず遠隔対話での上手く褒めるために重要な行動を明らかにする報告ができると考えている。

本研究での取り組みにより、上手く褒めるために重要な行動が明らかになること、上手く褒めるための能力の向上や褒め方の上手さを評価するシステムの開発が期待される。どのように褒めるとよいか悩んでいる方も多くいると考えられるが、様々な場面、状況、関係性において上手く褒めるために重要な行動を明らかにし、ひとりでも多くの方の悩みを解決できるようにしていきたい。さらに、褒める際の言語・非言語行動を評価し、「あなたの褒め方は60点です。もっと口角をあげて、相手のエピソードの内容を含めて褒めると良い関係性が築けますよ」のようなフィードバックを行う評価システムの開発を行い、ひとりでも多くの方の上手く褒めるための能力を向上させていきたい。

## 参考文献

- [1] T.M. Kalis, K.J. Vannest, and R. Parker. Praise counts: Using self-monitoring to increase effective teaching practices. *Preventing School Failure: Alternative Education for Children and Youth*, Vol. 51, No. 3, pp. 20–27, 2007.
- [2] L.N. Jenkins, M.T. Floress, and W. Reinke. Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, Vol. 52, No. 5, pp. 463–476, 2015.
- [3] J. Henderlong and M.R. Lepper. The effects of praise on children ’ s intrinsic motivation: A review and synthesis. *Psychological Bulletin*, Vol. 128, No. 5, pp. 774–795, 2002.
- [4] 大西俊輝, 柴田万里那, 呉健朗, 石井亮, 富田準二, 宮田章裕. 対話における上手い褒め方のモデリングの基礎検討. 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル (DICOMO2019) , 第 2019 卷, pp. 656–662, 2019.
- [5] 大西俊輝, 柴田万里那, 山内愛里沙, 呉健朗, 石井亮, 富田準二, 宮田章裕. 褒め方の上手さの推定における頭部・顔部の効果. 情報処理学会グループウェアとネットワークサービスワークショップ 2019 論文集, 第 2019 卷, pp. 1–6, 2019.
- [6] Toshiki Onishi, Arisa Yamauchi, Ryo Ishii, Yushi Aono, and Akihiro Miyata. Analyzing nonverbal behaviors along with praising. In *Proceedings of the 22nd ACM International Conference on Multimodal Interaction (ICMI’20)*, pp. 609–613, 2020.
- [7] 大西俊輝, 山内愛里沙, 大串旭, 石井亮, 青野裕司, 宮田章裕. 褒める行為における頭部・顔部の振舞いの分析. 情報処理学会論文誌, Vol. 62, No. 9, pp. 1620–1628, 2021.
- [8] J. Brophy. Teacher praise: A functional analysis. *Review of Educational Research*, Vol. 51, No. 1, pp. 5–32, 1981.
- [9] R.P. Ennis, D.J. Royer, K.L. Lane, H.M. Menzies, W.P. Oakes, and L.E. Schellman. Behavior-specific praise: An effective, efficient, low-intensity strategy to support student success. *Beyond Behavior*, Vol. 27, No. 3, pp. 134–139, 2018.

- 
- [10] R. Anderson, S.T. Manoogian, and J.S. Reznick. The undermining and enhancing of intrinsic motivation in preschool children. *Journal of personality and social psychology*, Vol. 34, No. 5, pp. 915–922, 1976.
- [11] D. Stipek. Study2: 2-5-year-olds’ reactions to success and failure and the effects of praise. *Monographs of the Society for Research in Child Development*, Vol. 57, pp. 39–59, 1992.
- [12] L. Batrinca, Mana, N., B. Lepri, N. Sebe, and F. Pianesi. Multimodal personality recognition in collaborative goal-oriented tasks. *IEEE Transactions on Multimedia*, Vol. 14, No. 4, pp. 659–673, 2016.
- [13] F. Valente, S. Kim, and P. Motlicek. Annotation and recognition of personality traits in spoken conversations from the ami meetings corpus. In *Proceedings of INTER-SPEECH*, pp. 1183–1186, 2012.
- [14] D. Jayagopi, D. Sanchez-Cortes, K. Otsuka, J. Yamato, and D. Gatica-Perez. Linking speaking and looking behavior patterns with group composition, perception, and performance. In *Proceedings of the 14th ACM international conference on Multimodal interaction (ICMI’12)*, pp. 433–440, 2012.
- [15] A.H. Anderson, M. Bader, E.G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, and C. Sotillo. The hrc map task corpus. *Language and speech*, Vol. 34, No. 4, pp. 351–366, 1991.
- [16] J. Carletta. Unleashing the killer corpus: experiences in creating the multi-everything ami meeting corpus. *Language Resources and Evaluation*, Vol. 41, No. 2, pp. 181–190, 2007.
- [17] J. Kickul and G. Neuman. Emergent leadership behaviors: The function of personality and cognitive ability in determining teamwork performance and ksas. *Journal of Business and Psychology*, Vol. 15, No. 1, pp. 27–51, 2000.
- [18] V. Ramanarayanan, C.W. Leong, L. Chen, G. Feng, and D. Suendermann-Oeft. Evaluating speech, face, emotion and body movement time-series features for automated multimodal presentation scoring. In *Proceedings of the 17th ACM on International Conference on Multimodal Interaction (ICMI’15)*, pp. 23–30, 2015.
- [19] L. Chen, G. Feng, J. Joe, C.W. Leong, C. Kitchen, and C.M. Lee. Towards automated assessment of public speaking skills using multimodal cues. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI’14)*, pp. 200–203, 2014.

- 
- [20] Y. Yagi, S. Okada, S. Shiobara, and S. Sugimura. Predicting multimodal presentation skills based on instance weighting domain adaptation. *Journal on Multimodal User Interfaces*, pp. 1–16, 2021.
- [21] S. Park, H.S. Shim, M. Chatterjee, K. Sagae, and L.P. Morency. Computational analysis of persuasiveness in social multimedia: A novel dataset and multimodal prediction approach. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI'14)*, pp. 50–57, 2014.
- [22] S. Okada, Y. Ohtake, Y.I. Nakano, Y. Hayashi, H.H. Huang, Y. Takase, and K. Nitta. Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI'16)*, pp. 169–176, 2016.
- [23] D. Sanchez-Cortes, O. Aran, M.S. Mast, and D. Gatica-Perez. A nonverbal behavior approach to identify emergent leaders in small groups. *IEEE Transactions Multimedia*, Vol. 14, No. 3, pp. 816–832, 2012.
- [24] L.S. Nguyen, D. Frauendorfer, M.S. Mast, and D. Gatica-Perez. Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior. *IEEE Transactions Multimedia*, Vol. 16, No. 4, pp. 1018–1031, 2014.
- [25] R. Ishii, K. Otsuka, S. Kumano, R. Higashinaka, and J. Tomita. Analyzing gaze behavior and dialogue act during turn-taking for estimating empathy skill level. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI'18)*, pp. 31–39, 2018.
- [26] Z.X. Tan, A. Goel, T.S. Nguyen, and D.C. Ong. A multimodal lstm for predicting listener empathic responses over time. In *OMG-Empathy Challenge workshop at the 14th IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, 2019.
- [27] M. Soleymani, K. Stefanov, H.S. Kang, J. Ondras, and J. Gratch. Multimodal analysis and estimation of intimate self-disclosure. In *Proceedings of the 21st ACM International Conference on Multimodal Interaction (ICMI'19)*, pp. 59–68, 2019.
- [28] P. Ekman and W.V. Friesen. Manual for the facial action coding system. *Palo Alto: Consulting Psychologists Press*, 1977.
- [29] F. Nihei, Y.I. Nakano, Y. Hayashi, H.H. Hung, and S. Okada. Predicting influential statements in group discussions using speech and head motion information. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI'14)*, pp. 136–143, 2014.

- 
- [30] M.A. McDaniel, D.L. Whetzel, F.L. Schmidt, and S.D. Maurer. The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of applied psychology*, Vol. 79, No. 4, pp. 599–616, 1994.
- [31] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt, M. Lhommet, and G. Lucas. Simsensei kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (AAMAS'14)*, pp. 1061–1068, 2014.
- [32] S.H. Kang and J. Gratch. Virtual humans elicit socially anxious interactants' verbal self-disclosure. *Computer Animation and Virtual Worlds*, Vol. 21, No. 3–4, pp. 473–482, 2010.
- [33] T.S. Polzin. Verbal and non-verbal cues in the communication of emotions. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2429–2432, 2000.
- [34] H. Brugman and A. Russel. Annotating multimedia / multi-modal resources with elan. In *Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC'04)*, pp. 2065–2068, 2004.
- [35] L. Huang, L.P. Morency, and J. Gratch. Parasocial consensus sampling: Combining multiple perspectives to learn virtual human behavior. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS '10)*, pp. 1265–1272, 2010.
- [36] P.E. Shrout and J.L. Fleiss. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, Vol. 86, No. 2, pp. 420–428, 1979.
- [37] P. Ekman. Movements with precise meanings. *Journal of communication*, Vol. 26, No. 3, pp. 14–26, 1976.
- [38] A. Kendon. Some functions of gaze-direction in social interaction. *Acta psychologica*, Vol. 26, pp. 22–63, 1967.
- [39] J.C. McCroskey and V.P. Richmond. Fundamentals of human communication: An interpersonal perspective. *Waveland PressInc*, pp. 169–214, 1996.
- [40] T. Baltrušaitis, P. Robinson, and L.P. Morency. Openface: An open source facial behavior analysis toolkit. In *IEEE Winter Conference on Applications of Computer Vision (WACV'16)*, pp. 1–10, 2016.
- [41] L. Breiman. Random forests. *Machine Learning*, Vol. 45, No. 1, pp. 5–32, 2001.

- 
- [42] J. Bergstra, D. Yamins, and D.D. Cox. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in Science Conferences (SciPy'13)*, pp. 13–20, 2013.
- [43] J.H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, Vol. 29, No. 5, pp. 1189–1232, 2001.
- [44] J.M. Brett. Culture and negotiation. *International journal of psychology*, Vol. 35, No. 2, pp. 97–104, 2000.
- [45] G. Doherty-Sneddon, A. Anderson, C. O'malley, S. Langton, S. Garrod, and V. Bruce. Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of experimental psychology: applied*, Vol. 3, No. 2, pp. 105–125, 1997.

# 研究業績

## 査読付き論文誌

- (1) Toshiki Onishi, Arisa Yamauchi, Asahi Ogushi, Ryo Ishii, Atsushi Fukayama, Takao Nakamura and Akihiro Miyata: Modeling Japanese Praising Behavior by Analyzing Audio and Visual Behaviors. *Frontiers in Computer Science* (2022年2月).
  - (2) 大西俊輝, 山内愛里沙, 大串旭, 石井亮, 青野裕司, 宮田章裕: 褒める行為における頭部・顔部の振舞いの分析, *情報処理学会論文誌*, Vol.62, No.9, pp.1620–1628 (2021年9月).
  - (3) 柴田万里那, 大西俊輝, 呉健朗, 宮田章裕: 柔らかい物体の動きによる共感表現方法の効果, *情報処理学会論文誌*, Vol.62, No.1, pp.26–34 (2021年1月).
  - (4) 立花巧樹, 呉健朗, 富永詩音, 大西俊輝, 鈴木颯馬, 諏訪博彦, 宮田章裕: 実世界オブジェクトを用いた生活空間内における事故予測支援手法, *情報処理学会論文誌*, Vol.62, No.1, pp.35–43 (2021年1月).
- 

## 査読付き国際会議

- (1) Kenro Go, Toshiki Onishi, Asahi Ogushi and Akihiro Miyata: Conversational Agents Replying with a Manzai-style Joke. *Proceedings of the 33rd Australian Conference on Human-computer-interaction (OzCHI '21)* (2021年11月).
  - (2) Toshiki Onishi, Arisa Yamauchi, Ryo Ishii, Yushi Aono and Akihiro Miyata: Analyzing Nonverbal Behaviors along with Praising. *Proceedings of the 22nd ACM International Conference on Multimodal Interaction (ICMI '20)*, pp.609–613 (2020年10月).
- 

## 査読付き国内会議

- (1) 呉健朗, 大西俊輝, 大串旭, 宮田章裕: ノリツッコミを行う対話型エージェント, *情報処理学会インタラクション2022 論文集* (2022年2月掲載予定).
- (2) 大串旭, 大西俊輝, 田原陽平, 石井亮, 深山篤, 中村高雄, 宮田章裕: 言語特徴に着目した褒め方の上手さの推定モデルの検討, *情報処理学会グループウェアとネットワークサービスワークショップ2021 論文集*, Vol.2021, pp.1–8 (2021年11月).
- (3) 大西俊輝, 柴田万里那, 山内愛里沙, 呉健朗, 石井亮, 富田準二, 宮田章裕: 褒め方の上手さの推定における頭部・顔部の効果, *情報処理学会グループウェアとネットワークサービスワークショップ2019 論文集*, Vol.2019, pp.1–6 (2019年11月).

- (4) 中原涼太, 長岡大二, 呉健朗, 大西俊輝, 柴田万里那, 宮田章裕: 複数対話型エージェントの役割分担によるユーモア生成システムの基礎検討, 情報処理学会グループウェアとネットワークサービスワークショップ2018 論文集, Vol.2018, pp.1-8 (2018年11月).

---

## 研究会・シンポジウム

- (1) 丸山葉, 大西俊輝, 大串旭, 呉健朗, 大澤正彦, 宮田章裕: 意見伝達を仲介する対話型エージェントの基礎検討, HAI シンポジウム 2022 (2022年3月掲載予定).
- (2) 得田舜介, 呉健朗, 田中柊羽, 大西俊輝, 大串旭, 宮田章裕: 英語でボケる発言を行うエージェントの基礎検討, 情報処理学会インタラクシオン 2022 論文集 (2022年2月掲載予定).
- (3) 大串旭, 大西俊輝, 山内愛里沙, 石井亮, 杵渕哲也, 青野裕司, 宮田章裕: 言葉づかいに着目した褒め方の上手さの推定モデルの基礎検討, 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル (DICOMO2021), Vol.2021, pp.791-797 (2021年7月).
- (4) 田中柊羽, 呉健朗, 大西俊輝, 大串旭, 武藤佑太, 宮田章裕: たとえツッコミを行う対話型エージェントの基礎検討, 情報処理学会インタラクシオン 2021 論文集, pp.639-641 (2021年3月).
- (5) 大串旭, 大西俊輝, 山内愛里沙, 石井亮, 杵渕哲也, 青野裕司, 宮田章裕: 上手く褒めるために効果的な言葉づかいの調査, 情報処理学会インタラクシオン 2021 論文集, pp.714-718 (2021年3月).
- (6) 山内愛里沙, 大西俊輝, 武藤佑太, 石井亮, 青野裕司, 宮田章裕: 音声および視線・表情・頭部運動に基づく上手い褒め方の評価システムの検討, 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル (DICOMO2020), Vol.2020, pp.98-106 (2020年7月).
- (7) 山内愛里沙, 大西俊輝, 呉健朗, 武藤佑太, 石井亮, 青野裕司, 宮田章裕: 表情・音声を用いた褒め方の上手さを評価するシステムの基礎検討, 情報処理学会インタラクシオン 2020 論文集, pp.247-252 (2020年3月).
- (8) 立花巧樹, 大西俊輝, 鈴木颯馬, 富永詩音, 呉健朗, 宮田章裕: 生活空間における危険予測支援システムの基礎検討, 情報処理学会グループウェアとネットワークサービスワークショップ 2019 論文集, Vol.2019, pp.99-102 (2019年11月).

- (9) 立花巧樹, 富永詩音, 大西俊輝, 呉健朗, 宮田章裕: ベビーカー利用時における周囲への動作予告システムの実装, 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル (DICOMO2019), Vol.2019, pp.1273–1279 (2019年7月).
  - (10) 大西俊輝, 柴田万里那, 呉健朗, 石井亮, 富田準二, 宮田章裕: 対話における上手い褒め方のモデリングの基礎検討, 情報処理学会シンポジウム論文集, マルチメディア、分散、協調とモバイル (DICOMO2019), Vol.2019, pp.656–662 (2019年7月).
  - (11) 柴田万里那, 大西俊輝, 呉健朗, 宮田章裕: 柔らかい物体の動きによる共感表現方法の基礎検証, 情報処理学会研究報告 (UBI), ユビキタスコンピューティングシステム, Vol.2019-UBI-62, No.11, pp.1–6 (2019年6月).
  - (12) 立花巧樹, 富永詩音, 大西俊輝, 呉健朗, 宮田章裕: ベビーカー利用時における周囲への動作予告手法の基礎検討, 情報処理学会インタラクシオン2019論文集, pp.179–181 (2019年3月).
  - (13) 柴田万里那, 大西俊輝, 呉健朗, 長岡大二, 中原涼太, 宮田章裕: 柔らかい物体の動きによる共感表現方法の基礎検討, 情報処理学会インタラクシオン2019論文集, pp.572–575 (2019年3月).
- 

## 受賞

- (1) 情報処理学会グループウェアとネットワークサービスワークショップ2021 ベストペーパー賞, 言語特徴に着目した褒め方の上手さの推定モデルの検討, 受賞者: 大串旭, 大西俊輝, 田原陽平, 石井亮, 深山篤, 中村高雄, 宮田章裕 (2021年11月).
- (2) マルチメディア、分散、協調とモバイル (DICOMO2020) シンポジウム 優秀論文賞, 音声および視線・表情・頭部運動に基づく上手い褒め方の評価システムの検討, 受賞者: 山内愛里沙, 大西俊輝, 武藤佑太, 石井亮, 青野裕司, 宮田章裕 (2020年7月).
- (3) 情報処理学会グループウェアとネットワークサービスワークショップ2019 ベストプレゼンテーション賞, 褒め方の上手さの推定における頭部・顔部の効果, 受賞者: 大西俊輝 (2019年11月).
- (4) マルチメディア、分散、協調とモバイル (DICOMO2019) シンポジウム ヤングリサーチ賞, 対話における上手い褒め方のモデリングの基礎検討, 受賞者: 大西俊輝 (2019年7月).
- (5) 情報処理学会グループウェアとネットワークサービスワークショップ2018 ベストペーパー賞, 複数対話型エージェントの役割分担によるユーモア生成システムの基礎検討, 受賞者: 中原涼太, 長岡大二, 呉健朗, 大西俊輝, 柴田万里那, 宮田章裕 (2018年11月).