

推薦論文

# 褒める行為における頭部・顔部の振舞いの分析

大西 俊輝<sup>1</sup> 山内 愛里沙<sup>1</sup> 大串 旭<sup>2</sup> 石井 亮<sup>3</sup> 青野 裕司<sup>3</sup> 宮田 章裕<sup>2,a)</sup>

受付日 2020年12月1日, 採録日 2021年6月7日

**概要:** 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が, 相手を上手く褒めるためには言語的・非言語的行動をどのように用いれば良いかは明らかにされていない。そこで我々は, 人間の言語的・非言語的行動を利用して, 対話における褒め方の上手さを分析する取り組みに着目する。本稿では, 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, 言語的・非言語的行動の中から頭部と顔部の振舞いの着目し, 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舞いが重要であるかを検証した。その結果, 頭部の向きや視線の方向に関する振舞い, 口角や口の開きに関する振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, 自分が相手を上手く褒められたかどうか判断する際は, 相手の瞼や瞬きの様子に注目すれば良い可能性が示唆された。

**キーワード:** マルチモーダルインタラクション, コミュニケーション, 褒める

## Analyzing Head and Face Behaviors along with Praising

TOSHIKI ONISHI<sup>1</sup> ARISA YAMAUCHI<sup>1</sup> ASAHI OGUSHI<sup>2</sup> RYO ISHII<sup>3</sup> YUSHI AONO<sup>3</sup> AKIHIRO MIYATA<sup>2,a)</sup>

Received: December 1, 2020, Accepted: June 7, 2021

**Abstract:** Praising behavior is considered to be a verbal and nonverbal behavior that expresses praise for the behavior and character of the target. However, how to use verbal and nonverbal behaviors to successfully praising is not clear. Therefore, we focus on attempts to analyze praising behavior in dialogue using verbal and nonverbal behaviors. In this paper, we focus on the behavior of the head and face and we investigate how to use the head and face behaviors to praise a partner. First, we created a dialogue corpus including head and face behaviors of the praiser and the receiver, as well as scores of the praising skills. Second, we constructed a machine learning model to estimate the evaluation of praise skills using features related to head and face behavior, and we evaluated the importance of head and face behavior of the praiser and receiver. As a result, we clarified that the behaviors related to the direction of the head and gaze, and the angle of the mouth and the opening of the mouth were important for successfully praising. Also, we suggested that we should pay attention to the eyelids and the blink of the eye when judging whether we were able to praise a partner.

**Keywords:** multimodal interaction, communication, praise

<sup>1</sup> 日本大学大学院総合基礎科学研究科  
Graduate School of Integrated Basic Sciences, Nihon University, Setagaya, Tokyo 156-8550, Japan

<sup>2</sup> 日本大学文理学部  
College of Humanities and Sciences, Nihon University, Setagaya, Tokyo 156-8550, Japan

<sup>3</sup> 日本電信電話株式会社 NTT 人間情報研究所  
NTT Human Informatics Laboratories, NTT Corporation, Yokosuka, Kanagawa 239-0847, Japan

a) miyata.akihiro@acm.org

### 1. はじめに

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

本稿の内容は 2019 年 11 月の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 で報告され, グループウェアとネットワークサービス研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である。

る [1], [2]. また, 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]. 褒める行為に関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, 褒めることの効果や影響を明らかにしている. しかし, 相手を上手く褒めるために言語的・非言語的行動をどのように用いれば良いかは明らかにされていない. このため, 日常生活の観点において, 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問題がある. また, 工学的な観点において, 人の褒め方の上手さを判定するシステムを構築することや, 上手い褒め方を行う対話型エージェントを実現することが困難であるという問題がある. これらの研究背景より, 相手を上手く褒めるためには, 言語的・非言語的行動をどのように用いれば良いか明らかにする必要がある.

そこで我々は, 人間の言語的・非言語的行動を利用して, 対話における褒め方の上手さを分析する取り組みに着目する. 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上手い褒め方を行う対話型エージェントの実現が期待される. 本稿では, 対話における褒め方の上手さと人間の行動の関係の分析として, 言語的・非言語的行動の中から非言語的行動である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. また, 先行研究 [4], [5], [6] の改善を行い, 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の推定を行う. 本稿の貢献は, 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.

## 2. 関連研究

### 2.1 褒める行為に関連する研究事例

褒める行為に関連する研究事例は数多く存在する. 多くの研究事例で, 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, 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事例は少なく, 大半の事例は教育シーンにおける褒める行為を扱っている. Brophy 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. 教師による褒めることは, (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 [7]. Henderlong 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 [3]. また, 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されている [8].

### 2.2 個人の性格や能力を推定する研究事例

人間の言語的・非言語的行動を利用して, 特定のタスクやシーンにおける行動・能力を分析する研究事例は数多く存在する. Batrinca らは, ヒューマンマシンインタラクションとヒューマンインタラクションの 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 [9]. Aran と Gatica-Perez は, グループディスカッションに参加している個人の性格特性を予測している [10]. Biel らは, vlogger の表情を用いて, Big Five の性格特性を推定している [11]. Lin と Lee は, ELEA コーパスにおける 10 種類の性格特性のスコアの予測性能を向上させるために, 対象話者とその文脈上の対局者の両方の発声行動をモデル化するフレームワークを提案している [12]. Pianesi らは, 音声と視覚的な特徴量を用いて, 会議環境における性格特性を自動認識する調査を行っている [13]. Valente らは, 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実験を行っている [14]. Jayagopi らは, 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 [15]. Wörtwein らは, プレゼンテーションのトレーニングのための対話型のバーチャルオーディエンスを提案している. 非言語的行動の自動評価と, プレゼンタの行動のマルチモーダルモデリングに着目している [16]. Ramanarayanan らは, 3 つの異なる特徴量を用いて, 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 [17]. Chen らは, 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 [18]. Park らは, 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, 話者の感情の部分的な事前知識を持つことが, 説得力のレベルをより良く予測すること

に寄与することを示している [19]. Nguyen らは、就職面接における応募者と面接官の非言語的な合図を用いて、面接での好感度を自動的に予測するためのフレームワークを提案している [20]. Sanchez-Cortes らは、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ることで、新たに形成されたグループの創発的リーダーシップを推測するフレームワークを提案している [21]. Okada らは、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 [22]. Ishii らは、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 [23]. Soleymani らは、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 [24]. Tan らは、OMG-Empathy Dataset において、聞き手の評価を予測するために、local attention レイヤでのマルチモーダル Long Short Term Memory モデルを提案している [25].

### 3. アプローチ

2.2 節で取り上げた研究事例では、特定のタスクやシーンにおける言語的・非言語的行動から、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。具体的には研究対象となる能力やパフォーマンスを特定のタスクやシーンの全体を通して観測される人間の行動を用いて推定や分析を行っている。しかし我々は、個人の褒め方の上手さを対話全体を通して分析するのではなく、個々のシーンにおける褒める行為に着目する。褒める行為は、個人内でも変異が大きいと考えられるため、全体的な個人の褒め方の上手さを推定するよりも、個々のシーンにおける褒める行為を扱う方が適切であると考えられる。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 [26]、言語的・非言語的行動の中から非言語的行動である頭部と顔部の振舞いに着目する。

上記をふまえ、本稿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、シーンごとの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、どのような行動が重要であるか明らかにする取り組みを行う。我々はこれまでに対話において褒められた人が相手の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の推定を行ってきた [4], [5], [6]。しかし、主観的な評価となり結果に偏りが出る可能性が考えられた。そのため、今回は複数の評価

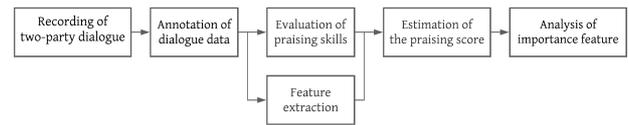


図 1 本研究のアプローチ

Fig. 1 Overall pipeline of our approach.

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の推定を行う。本稿では、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する。本研究でのアプローチを図 1 に示す。はじめに、褒める人と褒められる人の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、褒める人と褒められる人の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。

### 4. 対話コーパス

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。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

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 2・左のように参加者が互いに向き合って着座してもらった。このときの参加者間の距離は 180 cm とした。また、対話の収録は、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り、参加者の



図 2 2者対話の様子(左)と頭部のx軸, y軸, z軸方向の座標系(右)

Fig. 2 From left to right photograph of two-party dialogue (left) and coordinate system of the head (right).

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。また、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

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 [27] 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。発話シーンは、沈黙時間が400ミリ秒未満の連続した音声区間である。著者らで対話データをもとに議論を行い、発話の区切りが自然になるように沈黙時間が400ミリ秒未満の場合は1つの発話とした。次に、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。具体的には、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
- 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

本研究では、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。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。アノテータ間の Praise スコアの一

致率を評価するために、級内相関係数 (ICC) を利用した。Praise スコアの一致率は、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致率があり、信頼性のあるデータであることが示唆された。加えて、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次の判定・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。

### 5. 褒める行為における振舞いの分析

褒める人と褒められる人のどのような頭部と顔部の振舞いが、相手を上手く褒めるためには重要であるか明らかにするための分析を行う。具体的には、頭部と顔部の振舞い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。構築したモデルに基づいて、褒める人と褒められる人のどのような頭部と顔部の振舞いが、相手を上手く褒めるためには重要であるか分析する。

#### 5.1 特徴量抽出

顔画像処理ツールである OpenFace [28] を用いて、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、Action Units [29] に関する特徴量を抽出した。

**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2・右)、Praiseシーンの前後1秒ずつを含む範囲(図3)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パーセント値 ( $_{.p10}$ ), 90パーセント値 ( $_{.p90}$ ) 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

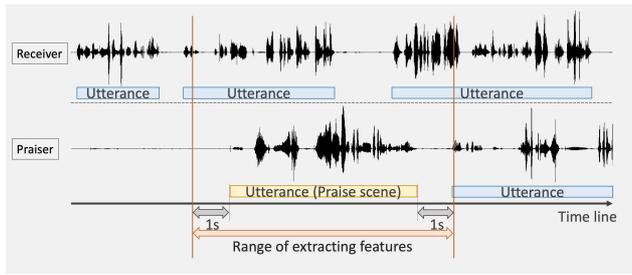


図 3 特徴量を抽出する範囲  
Fig. 3 The image of range of extracting features.

表 1 Action Units の内容  
Table 1 The list of Action Units.

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	上脣を上げる	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	脣を緊張させる	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		

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 パーセント値, 90 パーセント値を用いた。

**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 [29]. 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 1) の強度の分散, 中央値, 10 パーセント値, 90 パーセント値を用いた。

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。

### 5.2 褒め方の上手さの推定モデルの構築

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\*1。

\*1 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 2 各指標の平均値 (N = 100)  
Table 2 Evaluation results of model.

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

- 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 [30] を用いた。木の木数や木の深さなどのハイパパラメータは, Hyperopt [31] 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。

### 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 2 に示す。ベースライン (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ベースラインと提案するモデル, 提案するモデル間において対応のある t 検定を行った。その結果, ベースラインと Model P ( $t(99) = 6.541, p < .01$ ), Model R ( $t(99) = 3.342, p < .01$ ), Model P+R ( $t(99) = 8.696, p < .01$ ) であり, すべての組合せにおいて 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, Model P, Model R, Model P+R は, ベースラインよりも性能を向上させることができたと判断できる。また, Model P と Model R 間 ( $t(99) = 2.710, p < .01$ ), Model R と Model P+R 間 ( $t(99) = 4.900, p < .01$ ) において 1% 水準で有意差が認められ, Model P と Model P+R 間 ( $t(99) = 2.352, p < .05$ ) において 5% 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, 褒める人と褒められる人から抽

\* ベースラインとの間で  $p < .01$ 。

† Model R との間で  $p < .01$ 。

‡ Model P との間で  $p < .05$ 。

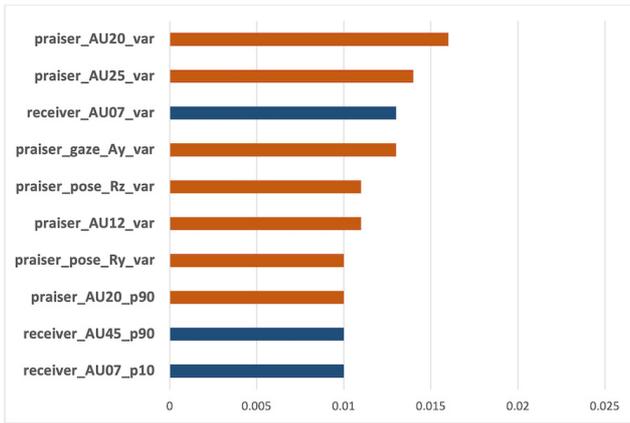


図 4 Model P+R における特徴量の重要度上位 10 件

Fig. 4 The top 10 features of importance in model P+R.

出した特徴量を用いた Model P+R は、褒める人から抽出した特徴量を用いた Model P と褒められる人から抽出した特徴量を用いた Model R よりも性能が高いと判断できる。以上より、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いえる。

#### 5.4 重要度の高い特徴量について

表 2 において最も推定精度が高かった Model P+R における特徴量の重要度上位 10 件を図 4 に示す。図 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。褒める人については、gaze\_Ay, pose\_Ry, 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\_Ry は頭部の y 軸回り（左右方向）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。これらの重要度の高かった特徴量について、Praise スコア高群と Praise スコア低群でどのような違いがあるのか分析を深める。表 3 に示すのは、高群・低群における各特徴量の中央値である。各スコアについて、群間で 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) が低群よりも高群で有意に大きいことが分かった。これはすなわち、高群においては、視線の上下、唇の両端の引き上げ、顎を下げずに唇を開く動作がより多く生じたことを意味している。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 に示す。

褒められる人については、AU07, AU45 に関連する特徴量の重要度が高かった。AU07 は顔を緊張させる動き、AU45 は瞬きをする動きを表している。自分が上手く褒め

表 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

Table 3 The median of the important features in the high and low praising score groups.

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**	0.436	0.248
receiver_AU07_var	0.161	0.154
praiser_gaze_Ay_var*	0.002	0.001
praiser_pose_Rz_var	0.001	0.000
praiser_AU12_var*	0.204	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

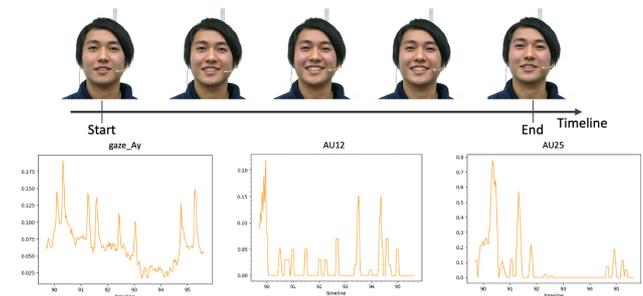


図 5 実際の対話シーンの画像と特徴量の値の一例

Fig. 5 The image of a dialogue scene and features values.

られたかどうか判断する際は、相手の顔や瞬きの様子に注目すればよい可能性がある。

#### 6. おわりに

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、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する取り組みを行った。具体的には、話者の頭部と顔部の振舞い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した。また、褒める人と褒められる人のどのような頭部と顔部の振舞いが相手を上手く褒めるために重要であるか分析した。その結果、頭部の向きや視線の方向に関する振舞い、口角や口の開きに関する振舞いを用いることで相手を上手く褒めることができる可能性が示唆された。また、自分が相手を上手く褒められたかどうか判断する際は、相手の顔や瞬きの様子に注目すれば良い可能性が示唆された。

本稿では上述の知見が得られたが、制約もある。はじめに、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た。そのため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である。しかし、褒められる

\*\* 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  $p < .05$ .

人の振舞いは相手に褒められた後に表れることも考えられる。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、推定精度が向上するか検証やどのような振舞いが重要であるか分析を行いたい。次に、本研究は人の褒め方の上手さを判定するシステムの構築や上手い褒め方を行う対話型エージェントの実現を目指しているが、現時点では褒めているシーンであるか否かの判定は行っていない。今後、本稿で得られた知見も生かし、褒めているシーンであるか否かの判定を自動的に行えるようにしていきたい。最後に、本稿で利用しているモダリティが十分ではない。今回は頭部・顔部の振舞いについて分析したが、コミュニケーションにおいてはジェスチャや発話内容も重要な役割を果たしていると考えられる。現在、これらのモダリティが褒める行為においてどのような役割を果たすか分析する準備を進めている。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。今後は、各シーン・関係性に依りて、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。

#### 参考文献

- [1] Kalis, T., Vannest, K. and Parker, R.: 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] Jenkins, L., Floress, M. and Reinke, W.: Rates and Types of Teacher Praise: A Review and Future Directions, *Psychology in the Schools*, Vol.52, No.5, pp.463–476 (2015).
- [3] Henderlong, J. and Lepper, M.: 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), Vol.2019, 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, Vol.2019, pp.1–6 (2019).
- [6] Onishi, T., Yamauchi, A., Ishii, R., Aono, Y. and Miyata, A.: Analyzing Nonverbal Behaviors along with Praising, *Proc. 22nd ACM International Conference on Multimodal Interaction (ICMI'20)*, pp.609–613 (2020).
- [7] Brophy, J.: Teacher praise: A fullcitional analysis, *Review of Educational Research*, Vol.51, No.1, pp.5–32 (1981).
- [8] Ennis, R., Royer, D., Lane, K., Menzies, H., Oakes, W. and Schellman, L.: Behavior-Specific Praise: An Effective, Efficient, Low-Intensity Strategy to Support Student Success, *Beyond Behavior*, Vol.27, No.3, pp.134–139 (2018).
- [9] Batrinca, L., Mana, N., Lepri, B., Sebe, N. and Pianesi, F.: Multimodal Personality Recognition in Collaborative Goal-Oriented Tasks, *IEEE Trans. Multimedia*, Vol.14, No.4, pp.659–673 (2016).
- [10] Aran, O. and Gatica-Perez, D.: One of a kind: Inferring personality impressions in meetings, *Proc. 15th ACM International Conference on Multimodal Interaction (ICMI'13)*, pp.11–18 (2013).
- [11] Biel, J., Teijeiro-Mosquera, L. and Gatica-Perez, D.: FaceTube: Predicting personality from facial expressions of emotion in online conversational video, *Proc. 14th ACM International Conference on Multimodal Interaction (ICMI'12)*, pp.53–56 (2012).
- [12] Lin, Y. and Lee, C.: Using Interlocutor-Modulated Attention BLSTM to Predict Personality Traits in Small Group Interaction, *Proc. 20th ACM International Conference on Multimodal Interaction (ICMI'18)*, pp.163–169 (2018).
- [13] Pianesi, F., Mana, N., Cappelletti, A., Lepri, B. and Zancanaro, M.: Multimodal recognition of personality traits in social interactions, *Proc. 10th International Conference on Multimodal Interfaces (ICMI'08)*, pp.53–60 (2008).
- [14] Valente, F., Kim, S. and Motlicek, P.: Annotation and Recognition of Personality Traits in Spoken Conversations from the AMI Meetings Corpus, *Proc. INTERSPEECH*, pp.1183–1186 (2012).
- [15] Jayagopi, D., Sanchez-Cortes, D., Otsuka, K., Yamato, J. and Gatica-Perez, D.: Linking speaking and looking behavior patterns with group composition, perception, and performance, *Proc. 14th ACM International Conference on Multimodal Interaction (ICMI'12)*, pp.433–440 (2012).
- [16] Wörtwein, T., Chollet, M., Schauerte, B., Morency, L., Stiefelhagen, R. and Scherer, S.: Multimodal Public Speaking Performance Assessment, *Proc. 17th ACM International Conference on Multimodal Interaction (ICMI'15)*, pp.43–50 (2015).
- [17] Ramanarayanan, V., Leong, C., Chen, L., Feng, G. and Suendermann-Oeft, D.: Evaluating Speech, Face, Emotion and Body Movement Time-series Features for Automated Multimodal Presentation Scoring, *Proc. 17th ACM International Conference on Multimodal Interaction (ICMI'15)*, pp.23–30 (2015).
- [18] Chen, L., Feng, G., Joe, J., Leong, C., Kitchen, C. and Lee, C.: Towards Automated Assessment of Public Speaking Skills Using Multimodal Cues, *Proc. 16th International Conference on Multimodal Interaction (ICMI'14)*, pp.200–203 (2014).
- [19] Park, S., Shim, H., Chatterjee, M., Sagae, K. and Morency, L.: Computational Analysis of Persuasiveness in Social Multimedia: A Novel Dataset and Multimodal Prediction Approach, *Proc. 16th International Conference on Multimodal Interaction (ICMI'14)*, pp.50–57 (2014).
- [20] Nguyen, L., Frauendorfer, D., Mast, M. and Gatica-Perez, D.: Hire me: Computational Inference of Hirability in Employment Interviews Based on Nonverbal Behavior, *IEEE Trans. Multimedia*, Vol.16, No.4, pp.1018–1031 (2014).
- [21] Sanchez-Cortes, D., Aran, O., Mast, M. and Gatica-Perez, D.: A Nonverbal Behavior Approach to Identify Emergent Leaders in Small Groups, *IEEE Trans. Mul-*

- timedia*, Vol.14, No.3, pp.816–832 (2012).
- [22] Okada, S., Ohtake, Y., Nakano, Y., Hayashi, Y., Huang, H., Takase, Y. and Nitta, K.: Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets, *Proc. 18th ACM International Conference on Multimodal Interaction (ICMI'16)*, pp.169–176 (2016).
- [23] Ishii, R., Otsuka, K., Kumano, S., Higashinaka, R. and Tomita, J.: Analyzing Gaze Behavior and Dialogue Act during Turn-taking for Estimating Empathy Skill Level, *Proc. 20th ACM International Conference on Multimodal Interaction (ICMI'18)*, pp.31–39 (2018).
- [24] Soleymani, M., Stefanov, K., Kang, H., Ondras, J. and Gratch, J.: Multimodal Analysis and Estimation of Intimate Self-Disclosure, *Proc. 21st ACM International Conference on Multimodal Interaction (ICMI'19)*, pp.59–68 (2019).
- [25] Tan, Z., Goel, A., Nguyen, T. and Ong, D.: A Multimodal LSTM for Predicting Listener Empathic Responses Over Time, *OMG-Empathy Challenge Workshop at the 14th IEEE International Conference on Automatic Face and Gesture Recognition (FG)* (2019).
- [26] Polzin, T.: Verbal and non-verbal cues in the communication of emotions, *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp.2429–2432 (2000).
- [27] Brugman, H. and Russel, A.: Annotating Multimedia/Multi-modal resources with ELAN, *Proc. 4th International Conference on Language Resources and Language Evaluation (LREC'04)*, pp.2065–2068 (2004).
- [28] Baltrušaitis, T., Robinson, P. and Morency, L.: OpenFace: An open source facial behavior analysis toolkit, *IEEE Winter Conference on Applications of Computer Vision (WACV'16)*, pp.1–10 (2016).
- [29] Ekman, P. and Friesen, W.: Manual for the Facial Action Coding System, *Palo Alto: Consulting Psychologists Press* (1977).
- [30] Breiman, L.: Random Forests, *Machine Learning*, Vol.45, No.1, pp.5–32 (2001).
- [31] Bergstra, J., Yamins, D. and Cox, D.: Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms, *Proc. 12th Python in Science Conferences (SciPy'13)*, pp.13–20 (2013).

#### 推薦文

本研究は、発話内容、声色、表情、姿勢、ジェスチャーといった多様なモダリティを用いた上手い褒め方をモデリングする手法の確立を目指したものであり、実用性・新規性の面から特筆すべきものであるため推薦する。

(グループウェアとネットワークサービス研究会主査 斉藤典明)



大西 俊輝 (学生会員)

2020年日本大学文理学部情報科学科卒業。同年日本大学大学院総合基礎科学研究科博士前期課程に進学、現在に至る。マルチモーダルインタラクション、ヒューマンコンピュータインタラクションの研究に従事。



山内 愛里沙 (学生会員)

2021年日本大学文理学部情報科学科卒業。同年日本大学大学院総合基礎科学研究科博士前期課程に進学、現在に至る。マルチモーダルインタラクション、ヒューマンコンピュータインタラクションの研究に従事。



大串 旭 (学生会員)

日本大学文理学部情報科学科に在学。マルチモーダルインタラクション、ヒューマンコンピュータインタラクションの研究に従事。



石井 亮

2008年東京農工大学大学院工学府情報工学専攻修士課程修了。同年日本電信電話株式会社入社。現在、NTT人間情報研究所特別研究員。2013年京都大学大学院情報学研究所博士後期課程修了。2011–2013年成蹊大学客員研究員。2019–2020年カーネギーメロン大学客員研究員。コミュニケーション科学、マルチモーダルインタラクションの研究に従事。ACM 15th Int. Conf. Multimodal Interaction (ICMI2014) Outstanding Paper Award, 他受賞。電子情報通信学会, 人工知能学会, ヒューマンインタフェース学会各会員。博士(情報学)。



青野 裕司 (正会員)

1999年大阪大学大学院基礎工学研究科博士後期課程修了。同年日本電信電話株式会社入社。NTTサイバースペース研究所, NTTアドバンステクノロジー, NTT研究企画部門, NTTメディアインテリジェンス研究所を経て, 現在, NTT人間情報研究所サイバネティックス研究プロジェクトマネージャー(主席研究員)。音声認識, 音声合成に関する研究に従事。電子情報通信学会各会員, 工学博士。



宮田 章裕 (正会員)

2005年日本電信電話株式会社入社。2008年慶應義塾大学大学院博士課程修了。2016年日本大学文理学部情報科学科准教授, 2021年より教授。HCI, アクセシビリティの研究に従事。情報処理学会2017年度・2018年度論文賞。ACM, 日本VR学会, HI学会, 日本DB学会各会員。博士(工学)。本会シニア会員。