

チュータリング対話の質評価に寄与する非言語情報の検討

辻本 海成^{1,a)} 角 康之^{1,b)}

概要：本稿は、大学生同士のピアチュータリング対話を題材とし、対話者の非言語行動から対話の質を評価する方法の可能性を検討する。大学生同士の学び合いの環境を良くしていくには、チューター（教える側）の技能を高め、それを継承していくことが重要である。したがって、チューター達は互いの技能を観察・評価し合い、真似すべき良い指導方法や、逆に避けるべき指導方法を指摘し、言語化していくことが重要である。そのためには、蓄積されたチュータリング対話のデータから、指導方法として良いシーンや悪いシーンを特定し、それらの理由を説明する原理を導き出すことが望まれる。本稿では、それらの手掛かりとして、チュータリングに参加している二者（チューターとチューティ）の非言語行動の量と時系列パターンに着目し、それらの組み合わせから対話シーンごとの質を数量化する方法を提案する。具体的には、計測された二者それぞれの発話と手作業、そしてそれらの共起に関する 14 の説明変数を使って、チュータリングのシーン毎の質を推測することを試みた。チュータリングデータからの注目すべきシーンの抽出と、それぞれのシーンの質評価には、熟練チューターに協力してもらった。それらのデータを利用して、チュータリングの質評価に寄与する説明変数の検討を行うために、説明変数間の相関分析、主成分分析、そして、質を評価するための回帰分析を行った。その結果、回帰式の作成にはまだ十分な結果は得られていないが、チューティ（学ぶ側）の関与が多く観察されるシーンが高く評価される、というチューター達の暗黙的な理解との符合を示唆する結果が得られた。

1. はじめに

本研究は、1 対 1 で実施される大学生のピアチュータリングを題材とし、チュータリング内で観察できる対話者の非言語行動からチュータリング対話の質評価を目指す。ここでは、椿本らが構築した学習支援組織「メタ学習ラボ」[1]にて行われるチュータリングを対象として議論を進める。メタ学習ラボは、著者らの所属する公立はこだて未来大学に 2012 年から導入された学生を中心とした学習支援組織である。メタ学習ラボに所属する学生チューター（教える側）は、自身のチュータリングスキルアップのために、自身が実施したチュータリングの振り返りやチューター同士でお互いのチュータリングを観察・評価することが多い。もちろん、チュータリングの振り返りや他人のチュータリングを評価することは、メタ学習の観点からも評価者のチュータリング技術向上につながると考える。しかしながら、そのプロセスは現状チューター自身の経験などに頼られることが多い。本来ならば、チューターとして良い・悪いチュータリングはどのようなものなのかを客観的に説

明できることが望ましい。そこで本稿では、チュータリングに参加している二者（チューターとチューティ）の非言語行動の量と時系列パターンに着目し、それらの組み合わせからチュータリング対話の質を数量化する方法を提案する。ただし、メタ学習ラボのチュータリングは 60 分で実施されるものであり、チュータリング全体を一意に良い・悪いと評価することは難しい。このことを踏まえて本稿では、チュータリング対話から意味のあるシーンを抽出して、それらのシーンごとの質を数量化する方法を提案する。

本稿では、はじめに、対象とするチュータリングや非言語情報などの基本アイデアを述べる。つぎに、当事者であるチューターのスキル向上を図りつつ、チュータリング対話から意味のあるシーンを得る試みである、チュータリング振り返りワークショップについて述べる。最後に、ワークショップで得られたシーンや熟練チューターから得たそれぞれのシーンに対する評価データとチュータリング内で観察できる非言語情報から、チュータリング対話の質評価に寄与する要素の検討について述べる。

ただし、本研究は、チュータリング内で観察できる非言語情報からチュータリング対話の質評価を実施し、チューター間にある暗黙的なチュータリングに対する理解の言語化を目指したものであり、チューティ（教えられる側）に

¹ 公立はこだて未来大学
Future University Hakodate

^{a)} k-tsujimoto@sumilab.org

^{b)} sumi@acm.org

与える学習効果については扱わない。

2. 関連研究

2.1 非言語インタラクションに関する研究

会話の中で交わされるインタラクションに含まれる非言語情報を対象とした研究は、これまでもいくつかなされてきた。会話の中におけるジェスチャー研究を行っている人物として Kendon が知られている [2]。会話参加者に注目し、その人が会話に参加しているかや会話への参与度、会話の主導権の推定を試みたもの [3][4][5] がある。また、医者と患者の 2 者間の会話を対象とし、そこで行われるインタビューを要約するために、対話データから複数の対話コーパスを生成し、ボトムアップ的に分析する方法を提案したもの [6] もある。複数人会話における非言語情報やインタラクションに注目した研究 [7] や、その中でも視線に着目した研究もある [8][9][10]。また Matsumura らは、ジェスチャーの共起性パターンを用いて聞き手の共感を分析した [11]。このように、インタラクションにおける非言語情報の重要性は広く認知されていると考える。

複数人で行われる会話を対象とした研究もなされてきた。大塚らは、会話の構造の推論を行うための確率的な枠組みを、会話参加者の視線、頭部方向および発話の有無に基づいて提案した [12][13]。Inoue らは、リアルな人型ロボットと人との 1 対 1 会話において、聞き手の会話への参与度へ与える要因について分析し、結果として笑い、あいづちを打つこと、うなずきが参与度と関連することを示唆した [14]。

会話の優位性 (dominance) を推定する研究として参考にした研究は次の通りである。Rienks は会話の優位性 (dominance) を発話に関する非言語情報や話者が行った質問の数から推定するモデルを Support Vector Machine を用いて提案した [15]。この研究で得られた分類器での会話優位性の推定正解率は 75% であり、この研究で取り上げられた非言語情報の中で最もモデルに寄与した要素は floorgrab であった。floorgrab は沈黙を破る発話として Rienks によって定義されたものであり、本稿においても会話の優位性に大きく影響を与える floorgrab の考えを利用する。また Nakano らは、複数人会話の優位性を視線情報と発話交替パターンから推定する手法を提案した [16]。そこで Nakano らは他人から集めた視線の量、互いに注目した視線の量、発話量、沈黙を破る発話を行った量の 4 つのパラメータを説明変数とした回帰分析を行った。Nakano らが試みた会話優位性のモデル化を回帰分析から行う考えを本稿でも参考にする。ただし、本稿で扱うチュータリングにおいて、視線を集めたりお互いに注目することは少ないと考えたため、後述する視線情報を拡張した共同作業空間を用いて議論を行う。また本稿では、会話の優位性をチュータリングにおける会話への積極性と読み替えて考

る。なぜならば、チュータリングは 1 対 1 会話であり、会話の優位性をとることはその会話を主導していると考えることができ、会話への積極性が高いと考えられるからである。

2.2 非言語情報を取得し議論する試み

会話における非言語情報を取得する試みとして Sumi らは IMADE (Interaction Measurement, Analysis, and Design Environment) ルームを提案した [17]。IMADE ルームには非言語情報を含むインタラクション行動を計測するために、環境カメラやマイク、モーションキャプチャ、アイマーカーレコーダ、データ統合と閲覧用サーバが設置されていた。IMADE ルーム内においてインタラクションを計測する際は、被験者 (被インタラクション計測者) は、それぞれ自分の体にモーションキャプチャ用トラッカーやアイマーカーレコーダ、マイクを装着していた。

取得したビデオなどを分析するツールも過去にいくつか提案されてきた。Burr はビデオを閲覧しながらアノテーションする作業の効率化を目指した VACA と呼ばれるシステムを提案した [18]。また、Sumi らは先述の IMADE ルームと同時に会話の構造分析を行うソフトウェア環境 iCorpusStudio を提案した [17]。この iCorpusStudio の特徴は、映像・音声の閲覧やラベリング作業を可能にするだけでなく、非言語情報間の時間構造分析を行うためのラベル間演算や身体動作、視線移動などの数値データ間計算を容易にし、非言語構造パターン解釈のための仮説検証を支援する点である。

2.3 会話参加者のスキル推定に関する研究

会話参加者のスキルを推定する研究もこれまでにいくつかなされてきた。コミュニケーション能力に着目した研究としては、岡田ら [19] の研究が挙げられる。岡田らは、複数人対話における会話参加者が表出するマルチモーダル情報 (単語の品詞・発話ターン・韻律情報・頭部動作量) と人事採用経験者による評価をもとにコミュニケーション能力を推定するモデルを構築した。また石井らは、発話交替時の視線遷移に着目して複数人対話における会話参加者の共感スキル推定モデルを構築した [20]。この推定モデルは、会話における参与役割を考慮して構築されたものであり、視線パターン以外の非言語的特徴量を用いた推定モデルよりも推定制度が高いことが示されている。本稿は、会話参加者の一般的なコミュニケーション能力といったスキルではなく、1 対 1 のチュータリングを対象とし、チュータリングにおける 2 人のインタラクションからチュータリングの質を推定するモデル構築を試みる。

2.4 チュータリングの分析やモデル化に関する研究

稲葉らは、協調学習における学習者間のインタラクシ

ンと学習効果の関連性に注目し、インタラクションを抽象化して望ましいインタラクションと比較し学習過程の分析・評価を行うシステムを提案した [21]. この研究では、協調学習者の複雑なインタラクションを抽象化する方法を提案し、対話プロトコルとして扱う。さらに、学習者に特定の学習効果がある学習理論に基づいた典型的なインタラクションパターンと学習者間の対話プロトコルを比較できる方法を提案した。稲葉らは、学習者間のインタラクションを主に言語情報を用いて分析しているが、本稿ではより単純で量的に分析できる非言語情報を用いて分析する。さらに、稲葉らは学習理論から導いたインタラクションパターンをいわば正解データとして扱っているが、本稿は実際にチュータリングを実施している人の評価を正解データとして扱うボトムアップ的なアプローチを試みる。

Grafsgaard は、チュータリング内での非言語行動の分析やモデル化をマルチモーダルな観点から行った [22]. この研究では、オンラインで行われるチュータリングにおいて、チュータリングを受ける人（本稿でいうチューティのこと）の非言語行動に注目されていた。注目した非言語行動は、顔の表情や姿勢、ジェスチャーであった。本稿では、チュータリングという性質から発話に関するモダリティも重要であるという観点から、身体的なモダリティだけでなく発話も含めて議論を行う。

3. 基本アイデア

本研究では、1対1で行われるチュータリングを対象に、非言語情報の観点からチュータリングの質を評価する。しかし、チュータリングを評価する際、チュータリング全体を対象に1つの評価を与えることは難しいと考えられる。例えば、チュータリングの全体の印象はそこまで良くないが、ある部分は非常に良いチュータリングができていた場合、どのように評価するかは評価者に大きく依存すると考える。そこで、チュータリングの中で見られる良い・悪いシーンとそこで観測できる非言語情報をもとに、チュータリング評価の要因となる非言語情報を見つける。具体的なアプローチとして本稿では、チュータリングを意味のある（注目できる）シーンごとに区切り、そのシーンについてチュータリングスコアをつける。それぞれのスコア付けされたシーンから、チュータリングの質に影響を与える非言語的ふるまいといった要素を明らかにすることを目指す。

3.1 チュータリングについて

チュータリングには様々な種類の形態が存在するが、本稿では1対1で行われるピア・チュータリング（本稿では、単にチュータリングと呼んでいる）を対象とする。特に、公立はこだて未来大学内に設置された学習支援組織メタ学習ラボに所属するチューター同士が実施したチュータリングロールプレイを本稿の分析対象とする。ここでいうチュー

タリングロールプレイは、学習支援組織メタ学習ラボ所属チューターの教育を目的とした模擬チュータリングのことである。本稿におけるチュータリングでは、教える側の人をチューター、教えられる側の人をチューティと呼ぶ。

3.2 着目する非言語情報

本稿で着目する非言語情報は、発話と作業空間への手の入り込み、頭部方向の3つである。発話から導き出せる発話交替という現象は、複数人会話の中で、会話参加者の会話への主導権を推定する際によく利用されるモダリティである。さらに、本稿では発話交替と同様に、チュータリングという場面においては、2人の作業空間への手の入り込みもチュータリングの評価するために重要であると考えられる。共同作業を行う複数人の間に生まれる空間に注目した研究は以前にもなされている。共同作業において Scott は、テーブル上の領域には personal, group, storage の3つの領域が存在することを提唱した [23]. Scott は、3つの領域をうまく活用しながら共同作業者はお互いのインタラクションを構築していることを示唆した。ここでは、主に共同作業を行う場であるとされている group 領域に着目して、発話（会話空間へのアクセス）と同様に作業（作業空間へのアクセス）についても考慮する。そこから、作業に対する主導権や積極性も表せると考える。ここで扱うチュータリングにおいて作業空間は、チュータリングに利用する教材や、ノート・ホワイトボード等の書き込むものの周辺であるとする。さらに、ただ作業空間への手の入り込みに着目するのではなく、チューターとチューティの注視方向についても考える。具体的には、2人が作業空間へ共同注視している場合（これを共同作業空間と呼ぶことにする）について着目する。つまり、共同作業空間とは、チューターとチューティの共同注視が行われている作業空間を指す（図1）。



図1 作業空間と共同作業空間

4. チュータリングデータ収集

ここでは、チュータリングビデオの収集からチュータリングをシーンごとに区切り、そのシーンについてスコア付けを行うまでの手法について述べる。さらに、チュータリ

ングビデオに付与した非言語情報に基づいたラベルデータについても説明する。

4.1 チュータリングビデオの収集

チュータリングビデオの撮影は、図 2 に示した環境で行う。撮影に使用するカメラは GoPro HERO7 Black とするが、特にカメラの画角を広角とせず通常の画角で撮影する。また音声の録音のために IC レコーダー (SONY ICD-UX560F) をチューターとチューティそれぞれに身に着けてもらう。

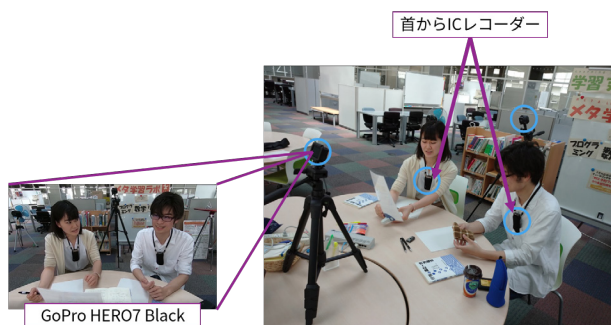


図 2 データ計測環境

本稿では、公立はこだて未来大学内に設置された学習支援組織メタ学習ラボに所属するチューターが実施したチュータリングロールプレイを対象にチュータリング動画の撮影を行う。各チュータリングセッションの長さは、メタ学習ラボで通常行われているセッションの時間と同じ 60 分とする。収集するチュータリングビデオは 4 本とし、チュータリングのトピックは数学とプログラミング 2 つずつとした。

4.2 シーン抽出を目的としたチュータリングワークショップ

チュータリングをシーンごとに区切るために、チュータリング振り返りワークショップを実施する。このワークショップは、実際にチュータリングを行っているチューターに参加してもらい、チューターのチュータリング技術向上を図りつつ、注目できるチュータリングシーンを得ようとする試みである。具体的には、撮影したチュータリングビデオに対してシーンの頭出しを行い、シーンとそれに対するコメントを書き込み、チューター間で共有できる取り組みである。この取り組みによってチューターには頭出しされたチュータリングシーンとそれに対するコメントが得られ、チュータリング事例集として活用できると考える。また、研究実施者は、実際にチュータリングを実施するチューターによって区切られたチュータリングシーンが得られる。

収集したチュータリングビデオを筆者が個人的に区切って評価を行う方法に比べて、チュータリングに対する専門

性を持った複数のチューターが関わることで、よりチュータリングの評価に適したシーンの抽出が可能であると考ええる。さらに、ただ機械的にチュータリングのシーン抽出を行う方法に比べて、チューターの技術向上を目指したチュータリングワークショップを実施する方法のほうが、研究実施者とチューターの双方にメリットがあり持続可能なデータ収集ができる可能性があると考えられる。以下に今回実施したチュータリング振り返りワークショップの概要を述べる。

チュータリング振り返りワークショップは、メタ学習ラボのチューターに集まってもらい、チュータリングビデオに対して頭出しやコメントができるプラットフォームを利用してもらう場とした。チュータリングの振り返りができるプラットフォームは、チャットサービスと動画共有サービスを利用した、チュータリング動画に対してコメントができるものである。このプラットフォームの目的は、撮影されたチュータリングビデオにおいて気になったシーンの頭出しを行いコメントをすることで、チュータリング事例集の作成をすることである。

具体的な利用例を以下に述べる。

- (1) 動画共有サービスにあるチュータリング動画を閲覧する。
- (2) チュータリングビデオの中で気になるシーンの頭出しを行いチャットサービスへ共有する (図 3 左)。
- (3) 頭出しされたチュータリングシーンの投稿に対してコメントをする (図 3 右)。
- (4) 他の人が頭出ししたシーンへのコメントや、自身のコメントへの返信を行い議論をする (図 3 右)。
- (5) 上記を繰り返す。



図 3 チュータリングシーンの頭出しとそれに対するコメントの例

今回実施したワークショップには 10 人のメタ学習ラボチューターが参加した。振り返りの対象としたチュータリングセッションは前節で述べた 4 本であった。また、ワークショップの全体の長さは 120 分とし、内訳はシーンの頭出し (40 分)・頭出しされたシーンへのコメント (50 分)・

ワークショップに関するアンケート回答（30分）であった。

このワークショップで得られるチュータリングシーンへのスコア付けは、アンケート形式で収集する。アンケートの構成は、ある一定の長さで区切られたチュータリングビデオに対して質問を1つとした。各ビデオに対して、このチュータリングを他のチューターへ勧めるかどうかを5段階のリッカート尺度で質問する。ここで得られたリッカート尺度の値（5に近い方が良いと判断された高いスコア）をそのシーンのスコアとみなす。ただし、チューター間での評価が分かれるシーンはチュータリングシーンのデータセットから除くことにする。つまり、本稿で用いるチュータリングシーンはチューター間評価の分散が小さいものを利用する。

4.3 作成するラベルデータ

本稿では、チュータリングにおける非言語情報をハンドラベリングで収集する。最初に、チュータリング内で観察できる基本的な非言語情報をラベリングする。今回作成するラベルは以下のとおりである。

- チューターとチューティの発話ラベル
- チューターとチューティの作業ラベル
- チューターとチューティの注視方向ラベル

なお、注視方向ラベルについては、作業空間を注視している場合と相手の顔を注視している場合を記録する。

さらに、Sumi らが提案した iCorpusStudio[17] を用いて、上記ラベル同士の演算を行う。本稿では以下の非言語情報やラベルをラベル同士の演算から得る。

- 共同作業空間ラベル
- チューターとチューティが共同作業空間で作業をしているラベル
- チューターとチューティが沈黙（1000ms）を破った回数
- チューターとチューティが発話と作業を同時に行っているラベル
- チューターがチューティの顔を見た回数
- 発話交替の回数
- チューターとチューティの平均発話長
- チューターとチューティの平均作業長

5. チュータリングワークショップの結果

実施したチュータリングワークショップの結果とそこで得られたチュータリングシーンについて述べる。今回ワークショップに参加したメタ学習ラボのチューターは10人であった。ワークショップ内で頭出しされたシーンは21シーンであり、シーンへのコメント投稿数は54であった。ここで用意したチュータリングビデオは4本の間に頭出しされたシーンの数に大きな違いは見られなかった。つまり、各チュータリングビデオそれぞれに、ある程度頭出し



図4 ワークショップ実施の様子

されたシーンが存在したということである。

このチュータリングワークショップは、ただ機械的にチュータリングのシーンを抽出するのではなく、チューターのチュータリングスキル向上を図りつつ、注目できるチュータリングシーンを得ようとする試みであった。そこで、このワークショップに参加したチューターの印象の概要を以下に述べる。はじめに、この試みがチュータリングスキル向上につながりそうか質問したところ、図5に示した通りおおむね良い評価が得られた。次に、このワークショップの所感を質問した。参加したチューターからは、「単純に他のチューターの考えを知ることができる機会が貴重で、さらにそれが目に見える形で残ることが良い」という意見や「チャットサービスを使うため議論に時間的制約がないことがじっくりコメントができる環境を作り出せている」といった意見が得られた。しかし、「頭出しされたシーンだけを見てもチュータリングの流れがわかりにくい」という意見や「対面での議論で感じる緊張感や手軽さが、文章で表現しないといけないこの試みでは薄い気がした」といった意見も得られた。これらの結果から、ただ機械的にチュータリングのシーン抽出を行う方法に比べて、チューターのチュータリングスキル向上を図りながら注目できるチュータリングシーンの抽出ができたと考える。

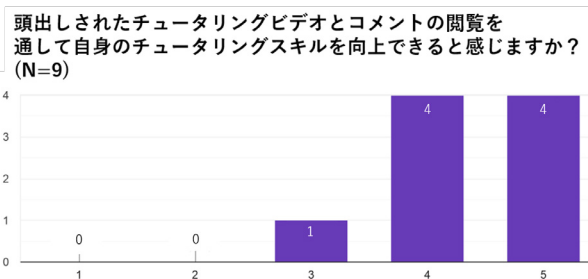


図5 ワークショップに対する印象を5段階のリッカート尺度で質問した結果

得られたチュータリングシーン 21 シーンの中から、シーンの重複を除いた 15 シーンを選び、後日、メタ学習ラボのチューターにスコアをつけてもらった。前章で述べた通り、アンケート形式でメタ学習ラボのチューター 6 人にスコア付けを行ってもらった。さらに、チューター間で評価が分かれるチュータリングシーンは分析の対象外とした。その結果、10 シーンのチュータリングシーンが得られた。また、それぞれのシーンの長さは 1 分程度から 2 分程度であった。

6. 非言語情報とチュータリング評価

6.1 各非言語情報とチュータリングスコアの間に見られる相関

チュータリングの中で観察できる非言語情報とチュータリングスコアそれぞれの関連性を明らかにするために、チュータリングスコアと 4.3 節で示した非言語情報の相関分析を行った。チュータリングスコアについては 1 から 5 までの整数で表された値、各非言語情報についてはそれぞれの値をチュータリングシーンの長さで割りチュータリングシーン間で正規化したものをそれぞれ利用した。相関分析の結果は図 6 に示した。図 6 の中で赤色に塗られた部分は負の相関が認められる値 ($r < -0.4$)、緑色に塗られた部分は正の相関が認められる値 ($r > 0.4$) を示した。

はじめに、チュータリングの評価に良い影響を与えると考えられる要素について議論する。チュータリングスコアと正の相関が認められた要素は、チューティの作業量・チューティが沈黙を破る発話の回数・チューティが発話と作業を同時に行う量・発話交替の回数の 4 つであった。これらから、チューティが多く会話の主導権を確保することやチューティ自身の作業が多いことが、チュータリングへ良い影響を与えられると考えられる。つまり、チューティが主導的なチュータリングシーンに対して良いチュータリングであると評価されることが示唆された。このことは、チュータリングが学びの場であり、課題を解決したい側（つまり、チューティ）がチュータリング内で積極性を望まれていることと直感的に合う結果が得られたと考える。

また、発話交替の回数にスコアと正の相関が認められたことについて述べる。発話交替の回数と、チューティの作業量・チューティの発話と作業を同時に行う量に正の相関が認められ、チューターの作業量・チューターがチューティの顔を見た回数に負の相関が認められた。これらから、発話交替が多く行われるシーンはチューティが教材に対して作業を行うシーンが多いと考えられる。つまり、発話交替が多いシーンはチューティ主導となっている場合が多いためチュータリングのスコアが高くなると考える。

つぎに、チュータリングの評価に悪い影響を与えると考えられる要素について議論する。チュータリングスコアと負の相関が認められた要素は、チューターが沈黙を破

る発話の回数であった。このことは、沈黙が存在する（発話をするチャンスがある）にもかかわらず、その沈黙をチューターが破り、会話の主導権を握るシーンが見られると、悪いチュータリングと評価されると考えられる。つまり、チューターが主導的なチュータリングシーンに対して悪いチュータリングであると評価されることが示唆された。

6.2 チュータリング評価モデルの作成

6.1 節で、チュータリングスコアと各非言語情報それぞれの関連性を明らかにするために相関分析を行ったが、相関分析では各要素同士にある 1 対 1 の関連性しかわからない。ここでは、目的変数をチュータリングスコアとし、説明変数をチュータリング内でみられる非言語情報の要素としてチュータリング評価モデルとなる回帰式を得る試みについて述べる。回帰式を得るための手段として重回帰分析を検討したが、図 6 に示した相関分析の結果の通り、説明変数間にかなり相関が認められた。つまり、説明変数間に多重共線性がみられると考える。そこで本稿では、4.3 節で示した非言語情報で主成分分析を行った。

表 1 主成分の寄与率と累積寄与率（累積寄与率が 0.8 を超えた後の主成分は省略した）

	第 1 主成分	第 2 主成分	第 3 主成分
標準偏差	2.58	1.73	1.24
寄与率	0.477	0.214	0.109
累積寄与率	0.477	0.691	0.800

主成分分析を行った結果、チュータリング内で観察できる非言語情報を説明する主成分は 3 つ得られた（表 1）。ただし、今回利用する主成分は、説明率 0.1 以上の主成分とした。また各主成分と非言語情報との主成分負荷量を表 2 に示した。

得られた各主成分が何を表しているか、主成分負荷量から考察する。はじめに、第 1 主成分はチュータリングに対する積極性を表していると考ええる。第 1 主成分において、正の値をとる主成分負荷量はすべてチューターが行う動作に見られた。一方で、負の値をとる主成分負荷量はチューティが行う動作と発話交替の回数に見られた。このことから、第 1 主成分はチューターが行動するほど正の値をとり、チューティが行動するほど負の値をとると考えられる。つまり、この主成分はチュータリングにおける積極性を表し、チューターが積極的な場合は正の値、チューティが積極的な場合は負の値をとる主成分であると考ええる。

つぎに、第 2 主成分は 1 回に行う発話や作業の長さを表していると考ええる。第 2 主成分において、チューター平均発話長・チューター平均作業長・チューティ平均発話長に顕著な値の主成分負荷量が見られた。つまり、この主成分は 1 回に行う発話や作業が長いほど負の値をとると考えられる。

	スコア	チューター 発話量	チューティ 発話量	チューター 作業量	チューティ 作業量	チューター沈黙を 破る発話回数	チューティ沈黙を 破る発話回数	チューター 作業発話同期	チューティ 作業発話同期	チューターが相手の 顔を見た回数	発話交替の 回数	チューター 平均発話長	チューター 平均作業長	チューティ 平均発話長	チューティ 平均作業長
スコア		-0.349	0.333	-0.279	0.427	-0.412	0.475	-0.395	0.429	0.025	0.511	0.192	0.226	0.207	0.282
チューター発話量	-0.349		-0.886	0.723	-0.631	0.328	-0.549	0.898	-0.665	0.257	-0.279	0.302	-0.032	-0.477	-0.712
チューティ発話量	0.333	-0.886		-0.480	0.459	-0.269	0.593	-0.761	0.503	-0.218	0.194	-0.010	0.333	0.763	0.512
チューター作業量	-0.279	0.723	-0.480		-0.727	0.242	-0.637	0.888	-0.730	0.396	-0.524	0.345	0.512	0.128	-0.573
チューティ作業量	0.427	-0.631	0.459	-0.727		-0.296	0.839	-0.616	0.976	-0.733	0.711	-0.350	-0.308	-0.073	0.674
チューター沈黙を 破る発話回数	-0.412	0.328	-0.269	0.242	-0.296		-0.193	0.379	-0.212	0.130	-0.314	-0.090	-0.160	-0.239	-0.280
チューティ沈黙を 破る発話回数	0.475	-0.549	0.593	-0.637	0.839	-0.193		-0.606	0.805	-0.656	0.712	0.019	-0.131	0.166	0.485
チューター 作業発話同期	-0.395	0.898	-0.761	0.888	-0.616	0.379	-0.606		-0.642	0.265	-0.468	0.153	0.101	-0.285	-0.543
チューティ 作業発話同期	0.429	-0.665	0.503	-0.730	0.976	-0.212	0.805	-0.642		-0.680	0.711	-0.400	-0.260	-0.056	0.605
チューターが相手の 顔を見た回数	0.025	0.257	-0.218	0.396	-0.733	0.130	-0.656	0.265	-0.680		-0.626	0.268	0.285	0.161	-0.242
発話交替の回数	0.511	-0.279	0.194	-0.524	0.711	-0.314	0.712	-0.468	0.711	-0.626		-0.107	-0.187	-0.241	0.044
チューターの 平均発話長	0.192	0.302	-0.010	0.345	-0.350	-0.090	0.019	0.153	-0.400	0.268	-0.107		0.637	0.431	-0.260
チューターの 平均作業長	0.226	-0.032	0.333	0.512	-0.308	-0.160	-0.131	0.101	-0.260	0.285	-0.187	0.637		0.782	-0.192
チューティの 平均発話長	0.207	-0.477	0.763	0.128	-0.073	-0.239	0.166	-0.285	-0.056	0.161	-0.241	0.431	0.782		0.229
チューティの 平均作業長	0.282	-0.712	0.512	-0.573	0.674	-0.280	0.485	-0.543	0.605	-0.242	0.044	-0.260	-0.192	0.229	

正の相関が認められる値 負の相関が認められる値

図 6 相関分析の結果

表 2 主成分負荷量

	第 1 主成分	第 2 主成分	第 3 主成分
チューターの発話量	0.839	0.380	0.345
チューティの発話量	-0.692	-0.645	-0.120
チューターの作業量	0.858	-0.181	0.230
チューティの作業量	-0.930	0.209	0.150
チューターによる 沈黙を破る発話回数	0.364	0.285	-0.178
チューティによる 沈黙を破る発話回数	-0.843	-0.023	0.370
チューターの 作業発話同期	0.848	0.247	0.189
チューティの 作業発話同期	-0.922	0.189	0.114
チューターが相手の 顔を見た回数	0.626	-0.347	-0.428
発話交替の回数	-0.655	0.293	0.586
チューターの 平均発話長	0.326	-0.560	0.537
チューターの 平均作業長	0.244	-0.845	0.339
チューティの 平均発話長	-0.130	-0.964	-0.002
チューティの 平均作業長	-0.694	-0.142	-0.412

第 1 主成分の解釈 チューターが積極的な場合は正の値、
チューティが積極的な場合は負の値をとる
第 2 主成分の解釈 1 回の発話や作業が長いほど負の値をとる
第 3 主成分の解釈 チューティの会話への積極性を表す

最後に、第 3 主成分はチューターの発話が多い中でチューティが多く発話したかを表していると考えられる。第 3 主成分において、注目できる主成分負荷量はチューター発話量・チューティ発話量・チューティが沈黙を破る発話回数・発話交替の回数・チューター平均発話長と考えられる。チュー

ターの発話量とチューターの平均発話長が正の値でなおかつ、チューティの発話量が負の値であることが確認できる。また、チューティが沈黙を破る発話回数と発話交替の回数が正の値であることが確認できる。このことから、この主成分はチューター主導の会話の中でチューティが多く会話のターンを取ろうとしたかを表す主成分であると考えられる。つまり、第 3 主成分はチューティの会話への積極性を表すと考えられる。しかし、ここで着目した主成分負荷量は第 1 主成分や第 2 主成分における主成分負荷量に比べて絶対値が小さいため、そもそも影響が少ない主成分である。

表 3 主成分を説明変数とした回帰分析の結果

	係数	標準誤差	t 値	p 値
切片	3.72	0.125	29.74	0.000
第 1 主成分	-0.412	0.127	-3.25	0.002
第 2 主成分	-0.240	0.127	-1.90	0.066
第 3 主成分	0.261	0.127	2.06	0.046

表 4 回帰分析の統計データ

重決定 R^2	0.321
補正 R^2	0.269
観測数	43
有意 F	0.002

主成分分析を行った後、目的変数をチュータリングスコアとし説明変数を主成分分析で得られた 3 つの主成分とした回帰分析を行った。チュータリングのスコアを y 、第 1 主成分から第 3 主成分までを x_1 から x_3 とした場合、以下の回帰式

$$y = -0.412x_1 - 0.240x_2 + 0.261x_3 + 3.72$$

が得られた。回帰分析の結果は表 3 に示した。また、回

帰分析の統計データを表 4 に示した。この回帰分析の観測数は 43 であり決定係数 R^2 は 0.321, 有意 F は 0.002 であった。

得られた回帰式それぞれの係数に着目すると, 第 1 主成分と第 2 主成分は負の相関が, 第 3 主成分は正の相関がある。つまり, チューターの積極性と 1 回の発話や作業の短さに負の相関が, チューティの会話への積極性に正の相関がそれぞれチュータリングスコアとの間に見出される。

得られた式の中で注目すべき傾向について議論する。まず, 第 1 主成分とその係数に注目する。第 1 主成分はチュータリングに対する積極性を表す主成分であり, チューターが積極的な場合は正の値, チューティが積極的な場合は負の値をとる主成分であった。この第 1 主成分の係数は負の値であったため, チューターが積極的な場合はスコアが下がると考えることが可能である。つまり, チューティが積極的なチュータリングのほうがチュータリングスコアが高くなる傾向が示唆された。さらに, この傾向は 6.1 節で議論した内容とそこまで相違は無いように見られる。また第 1 主成分の係数は 1% 水準で有意差が確認できたため, 強い傾向であると考ええる。つぎに, 第 3 主成分とその係数に着目する。第 3 主成分はチューティの会話への積極性を表す主成分であった。この第 3 主成分の係数は正の値であったため, チューティが会話への積極性をみせるとスコアが上がることを示唆された。この結果は, 先ほど述べた第 1 主成分に関する傾向と同じ傾向であった。しかし, 得られた回帰式に対する決定係数 R^2 が低いことや, 第 2 主成分に対応する偏回帰係数が 0 である帰無仮説を棄却できていないことから, この回帰式でチュータリングのスコアと非言語情報の関連性をすべて表したとは考えにくい。ただし, チュータリング評価は非言語情報のみでなく言語情報も含めて行うものであることを鑑みると, 非言語情報のみで求めた回帰式の決定係数 R^2 が低いことはそこまでの外れではないと考える。

最後に, 本稿で行った回帰分析で当てはまりの良い結果が得られなかった原因について議論する。原因の 1 つ目として, チュータリングを説明する説明変数の不足が考えられる。本稿では 4.3 節で述べた要素を説明変数として主成分分析を行った。しかしながら, 主成分を説明変数とした回帰分析で得られた式では, チュータリングのスコアに影響を与える要素の説明が完全にはできなかった。今後, 新たな説明変数の候補を言語情報 (例えば, 発話の中であいづちは区別する) を含めて検討していく必要がある。原因の 2 つ目としては, 今回チュータリングワークショップで得たチュータリングシーンの内容に大きな差があったことである。本稿で対象としたチュータリングは, 導入・実装・まとめ 3 つのフェーズで構成されている。ここで問題となることは, 3 つのフェーズ間でチューターに期待されるふるまいが異なることである。例えば, 導入フェーズに

おいてチューターはチューティの課題に対する問題やその背景を聞き出してチュータリング全体の計画を立てる必要がある。そのため, 導入フェーズにおいてチューターはチューティに対して多くの質問を行い, メモを取るなどの作業をする。一方で, 実装フェーズにおいてチューターは, チューティの課題解決の補助をする立場になるため, チューターが積極的に発話や作業をしないことが求められる。本稿で分析に利用したデータセットのフェーズ別内訳は, 導入が 3 シーンで実装が 7 シーンであった。つまり, 同じ高スコアのチュータリングシーンであっても, チューターに求められるふるまいの違いから, チューターが積極的なシーンとそうでないシーンが混ざったと考えられる。今後は, 各フェーズごとに分析を行い, フェーズ間で得られる回帰式に差が生まれるかを確かめていきたい。

7. おわりに

本稿では, 非言語情報の観点からチュータリングの評価をするためにチュータリングの質 (スコア) に影響を与える非言語情報について議論した。まずチュータリングの中で注目できるシーンとそのシーンに対するスコアを得るためにチュータリングワークショップを行った。つぎに, チュータリングスコアとそのチュータリング内でみられる非言語情報との関連性を調べるために主成分分析と回帰分析を行った。主成分分析と回帰分析を行った結果, 高い決定係数をもった回帰式は得られなかったが, チューティに積極性が見られるチュータリングに高いスコアがつけられる可能性が示唆された。

今後は, より統計的に有意なチュータリング評価モデルを作成するために, チュータリングの質に影響を与える非言語的ふるまいや要素を検討する。その際, チュータリング全体の平均モデルを作るのではなく, チューターの役割に応じたフェーズに分割して検討することが重要であると考ええる。

次に, 非言語情報から作成したチュータリングの質評価式を実際のチュータリング現場で利用し, その評価式で得られたチュータリングスコアをチューターに評価してもらうことで, チュータリングの質評価式の妥当性を評価したい。また, Higuchi ら [24] が提案した振り返りシステムのように, チュータリングに影響を与える非言語情報を手掛かりとした, チュータリング振り返りシステムを実現したい。

謝辞 本研究を実施するにあたり, ご討論頂いた公立はこだて未来大学教授富永敦子先生と同准教授中村美智子先生に感謝する。また, チュータリングビデオや評価データの収集にご協力頂いたメタ学習ラボチューターのみなさまに感謝する。

参考文献

- [1] 椿本弥生, 大塚裕子, 高橋理沙, 美馬のゆり. 大学生を中心とした持続可能な学習支援組織の構築とピア・チュータリング実践. 日本教育工学会論文誌, Vol. 36, No. 3, pp. 313–325, 2012.
- [2] A. Kendon. *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- [3] Yukiko I. Nakano and Ryo Ishii. Estimating user’s engagement from eye-gaze behaviors in human-agent conversations. In *Proceedings of the 15th International Conference on Intelligent User Interfaces, IUI ’10*, pp. 139–148, New York, NY, USA, 2010. ACM.
- [4] Ryo Ishii and Yukiko I. Nakano. An empirical study of eye-gaze behaviors: Towards the estimation of conversational engagement in human-agent communication. In *Proceedings of the 2010 Workshop on Eye Gaze in Intelligent Human Machine Interaction, EGIHMI ’10*, pp. 33–40, New York, NY, USA, 2010. ACM.
- [5] Misato Yatsushiro, Naoya Ikeda, Yuki Hayashi, and Yukiko I. Nakano. A dominance estimation mechanism using eye-gaze and turn-taking information. In *Proceedings of the 6th Workshop on Eye Gaze in Intelligent Human Machine Interaction: Gaze in Multimodal Interaction*, GazeIn ’13, pp. 13–18, New York, NY, USA, 2013. ACM.
- [6] Kenji Mase, Yuichi Sawamoto, Yuichi Koyama, Tomio Suzuki, and Kimiko Katsuyama. Interaction pattern and motif mining method for doctor-patient multi-modal dialog analysis. In *Proceedings of the ICMI-MLMI ’09 Workshop on Multimodal Sensor-Based Systems and Mobile Phones for Social Computing*, ICMI-MLMI ’09, pp. 6:1–6:4, New York, NY, USA, 2009. ACM.
- [7] Justine Cassell, Yukiko I. Nakano, Timothy W. Bickmore, Candace L. Sidner, and Charles Rich. Non-verbal cues for discourse structure. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics, ACL ’01*, pp. 114–123, Stroudsburg, PA, USA, 2001. Association for Computational Linguistics.
- [8] Yukiko I. Nakano, Gabe Reinstein, Tom Stocky, and Justine Cassell. Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1, ACL ’03*, pp. 553–561, Stroudsburg, PA, USA, 2003. Association for Computational Linguistics.
- [9] Pernilla Qvarfordt and Shumin Zhai. Conversing with the user based on eye-gaze patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’05*, pp. 221–230, New York, NY, USA, 2005. ACM.
- [10] Roel Vertegaal, Robert Slagter, Gerrit van der Veer, Gerrit van der Veer, and Anton Nijholt. Eye gaze patterns in conversations: There is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’01*, pp. 301–308, New York, NY, USA, 2001. ACM.
- [11] Kohei Matsumura, Yasuyuki Sumi, and Mitsuki Sugiya. Analyzing listeners’ empathy by their nonverbal behaviors in bibliobattle. *Journal of Information Processing*, Vol. 25, pp. 361–365, 2017.
- [12] 大塚和弘, 竹前嘉修, 大和淳司, 村瀬洋. 複数人物の対面会話を対象としたマルコフ切替えモデルに基づく会話構造の確率的推論. 情報処理学会論文誌, Vol. 47, No. 7, pp. 2317–2334, jul 2006.
- [13] Kazuhiro Otsuka, Hiroshi Sawada, and Junji Yamato. Automatic inference of cross-modal nonverbal interactions in multiparty conversations: ”who responds to whom, when, and how?” from gaze, head gestures, and utterances. In *Proceedings of the 9th International Conference on Multimodal Interfaces (ICMI 2007)*, pp. 255–262. ACM, 2007.
- [14] Koji Inoue, Divesh Lala, Shizuka Nakamura, Katsuya Takanashi, and Tatsuya Kawahara. Annotation and analysis of listener’s engagement based on multi-modal behaviors. In *Proceedings of the Workshop on Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction, MA3HMI ’16*, pp. 25–32, New York, NY, USA, 2016. ACM.
- [15] Rutger Rienks and Dirk Heylen. Dominance detection in meetings using easily obtainable features. In Steve Renals and Samy Bengio, editors, *Machine Learning for Multimodal Interaction*, pp. 76–86, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
- [16] Yukiko Nakano and Yuki Fukuhara. Estimating conversational dominance in multiparty interaction. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI ’12*, pp. 77–84, New York, NY, USA, 2012. ACM.
- [17] Yasuyuki Sumi, Masaharu Yano, and Toyoaki Nishida. Analysis environment of conversational structure with nonverbal multimodal data. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction, ICMI-MLMI ’10*, pp. 44:1–44:4, New York, NY, USA, 2010. ACM.
- [18] Brandon Burr. Vaca: A tool for qualitative video analysis. In *CHI ’06 Extended Abstracts on Human Factors in Computing Systems, CHI EA ’06*, pp. 622–627, New York, NY, USA, 2006. ACM.
- [19] 岡田将吾, 松儀良広, 中野有紀子, 林佑樹, 黄宏軒, 高瀬裕, 新田克己. マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定. 人工知能学会論文誌, Vol. 31, No. 6, pp. AI30–E1–12, 2016.
- [20] 石井亮, 熊野史朗, 大塚和弘. 話者継続・交替時における参与役割に応じた視線行動に基づく共感スキルの推定. ヒューマンインタフェース学会論文誌, Vol. 20, No. 4, pp. 447–456, 2018.
- [21] 稲葉晶子, 大久保亮二, 池田満, 溝口理一郎. 協調学習におけるインタラクション分析支援システム. 情報処理学会論文誌, Vol. 44, No. 11, pp. 2617–2627, nov 2003.
- [22] Joseph F. Grafsgaard. Multimodal analysis and modeling of nonverbal behaviors during tutoring. In *Proceedings of the 16th International Conference on Multimodal Interaction, ICMI ’14*, pp. 404–408, New York, NY, USA, 2014. ACM.
- [23] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, CSCW ’04*, pp. 294–303, New York, NY, USA, 2004. ACM.
- [24] Keita Higuchi, Ryo Yonetani, and Yoichi Sato. Egoscanning: Quickly scanning first-person videos with egocentric elastic timelines. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI ’17*, pp. 6536–6546, New York, NY, USA, 2017. ACM.