

# 非言語情報に着目したチュータリングの質評価と チューター熟練度の関連性

辻本 海成<sup>1,a)</sup> 角 康之<sup>1,b)</sup>

**概要:** チュータリングは主観的なリフレクションや評価を行うことが多いが、非言語情報からもチュータリングの評価ができると考える。主観的に評価されたチュータリングのスコアを非言語情報の観点から説明することを目指すアプローチで、チュータリングの評価を目指す。そこで、チュータリングを行う学習支援組織に所属するチューターへアンケート調査を行い、あるチュータリングシーンについて主観的な評価をしてもらった。本稿では、アンケート調査で評価してもらったスコアとチュータリングシーンに見られた非言語情報との関連性を議論する。アンケート調査で得られた評価とチュータリング内に見られた非言語情報から、学習支援組織所属チューター間に共通した評価基準があること、いくつか評価に影響を与える非言語的要因があることが示唆された。

## 1. はじめに

人が行う会話は、知識伝達や意思疎通などの手段であり、最も盛んに行われているコミュニケーションの1つである。会話の中で我々は言語情報はもちろん、視線、ジェスチャ、うなずき、あいづちといった非言語情報によってさまざまな意図を伝える。我々が交わしている非言語情報には一定のパターン（アイコンタクトのタイミングや量、手の位置、発話のタイミングなど）がある。さらに、非言語情報をお互いに無意識のうちに利用し、会話の制御を行ったり、自身の気持ちを伝えたりする。知識伝達で生じるインタラクションが相手に与える印象はについて、福島らはテキストデータと音声データからプレゼンテーションが聴衆者へ与える印象の推定を行っている [1]。ただ、テキストデータや音声データだけでなく、非言語情報も相手に与える印象に大きく関わることは容易に想像できると考えられる。

知識伝達の形態として、近年ではピア・チュータリング（本研究では、単にチュータリングと呼んでいる）の試みが注目されつつある。椿本らは学生を中心とした学習支援組織とそのための空間である「メタ学習ラボ」を構築し、メタ学習を目的としたチュータリング活動を行った。その結果、学習者にはチュータリングによるメタ認知の促進効果や、学習意欲の向上効果が得られる可能性を示唆した [2]。メタ学習ラボでは、大学2年生以上の学生チューターが

チューティ（主に大学1年生）に対して1対1のチュータリングという形で学習支援を行っている。チューターは教える側の人、チューティは教えられる側の人のことを指す。

先述のメタ学習ラボではチュータリングを評価する客観的な手法はあまりない。現在メタ学習ラボではチュータリングの評価を、自身のリフレクションや他人からの評価によって行っている。もちろん、チュータリングのリフレクション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。しかし、それぞれの評価は評価者の主観的な評価となる。チュータリングの個性は非常に重要で尊重されるべきものであるが、客観的な共通した評価の枠組みがあることで、チューターの育成の省力化や効率化が望めると考える。また、中園らは高等教育機関で実施されるアクティブラーニングについて批判的検討を行っている [3]。このように、アクティブラーニングの一部であるチュータリングについて議論を行うことは有益であると考えられる。

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。具体的には、非言語情報に基づいたチュータリングスコア算出モデルを作成することを目指す。チュータリングスコア算出モデルを作成するために、あるチュータリングで見られた非言語情報からボトムアップ的にモデルを作成するのではなく、人が付けたスコアからトップダウン的にモデルを作成する手法を試みる。本稿では、その取り組みの一つとして、あるシーンに分割されたチュータリングビデオをチュータリングを専門の学習支援組織に所属するチュー

<sup>1</sup> 公立はこだて未来大学  
Future University Hakodate

a) k-tsujimoto@sumilab.org

b) sumi@acm.org

ターにスコア付けしてもらい、ビデオ内に見られた非言語情報との関連性やチューターが持つ評価の判断基準について議論する。

## 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 [4]。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加度、会話の主導権の推定を試みたもの [5][6][7] や医者と患者の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 [8] がある。

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 [9]。Nakano らは、会話優位性を自動推定するために、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つのパラメーターに着目したモデルを提案した [10]。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加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、うなずきが参加度と関連することを示唆した [11]。

### 2.2 チュータリングにおける非言語行動に着目した研究

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 [12]。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（本研究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。

### 2.3 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 [13]。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、アイマークレコーダ、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クレコーダ、マイクを装着していた。

Sumi らは 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 [13]。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。

## 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。具体的なアプローチとして本研究では、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チュータリングスコアを付けを行う。それぞれのスコア付けされたシーンから、良いチュータリング事例集の作成やチュータリングを行う人の間にある暗黙知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、本研究では1対1で行われるピア・チュータリング（本研究では、単にチュータリングと呼んでいる）を対象とする。本研究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

本研究で着目する非言語情報は、発話交代と作業空間への手の入り込み、頭部方向の3つである。発話交代は、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、本研究では発話交代と同様に、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、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案した [14]。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、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）。

## 4. チュータリングの印象評価に関する調査

ここでは、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チュータリングスコアを付けを行うために実施したアンケート調査について述べる。このアンケート



図 1 作業空間と共同作業空間

調査では、チュータリングを普段から行う人からチュータリングに対する印象をスコアにしてもらい、そのシーンで見られた非言語情報との関連性を見ることを目的とする。

#### 4.1 データ計測

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。撮影に利用するカメラは GoPro HERO4 とする。チュータリングを行う 2 名には、チュータリングを 20 分から 30 分ほどで行うように教示する。本稿で収集するチュータリングビデオは、特に学習支援組織に所属していない人が行うチュータリングである。またチュータリングの題材は自由なものとするが、ラップトップ PC を利用しないように教示する。

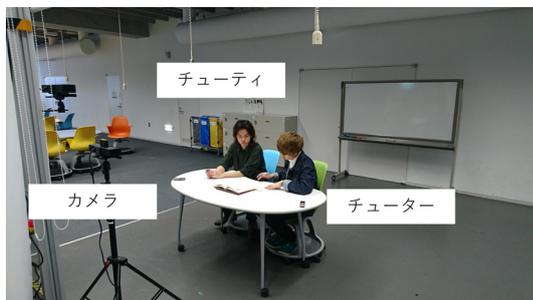


図 2 データ計測環境

#### 4.2 アンケート調査

撮影したチュータリングビデオから、4 つのシーンを切り出し、それぞれのシーンについて印象評価を行うためにアンケート調査を実施する。アンケートの対象は、ピア・チュータリングを普段から行う学習支援組織メタ学習ラボに所属するチューターと学習支援組織に所属経験がない一般学生とする。また、アンケートはオンラインで回答してもらい、回答に関する制約は設けない形とする。

#### 4.3 質問内容

今回作成したアンケートの質問内容について述べる。アンケートの構成は、ビデオ 1 本に対して質問が 3 つに加えて全体を通した質問 1 つとした。各ビデオに対して行った質問は以下の 3 つである。

- このチュータリングではどちらに会話の支配権があったと思いますか？
- このチュータリングをどの程度他人に勧めたいと思いますか？
- このチュータリングでは、チューティ（学習者）に自発的に学ぼうとする姿勢が感じられましたか？

それぞれの質問は 5 段階のリッカート尺度とし、1 つ目の会話の支配権に関する質問においては 1 をチューター、5 をチューティとした。他の 2 つ質問においては、1 を全く同意できない、5 を強く同意するとした。これと同じ質問を、合計 4 つのビデオに対して実施した。最後に、回答者がどのような観点で印象評価したのかを確認するために以下の質問を自由記述形式で行った。

- あなたが本アンケートの設問に対して判断する際に重視したことは何ですか？

#### 4.4 利用したチュータリングビデオ

今回アンケートで利用したチュータリングビデオは 4 本であった。それぞれのビデオは 3 分～5 分の長さとし、シーンの区切りは筆者が内容ベースで行った。つまり、筆者が主観的に 1 つのトピックについてチュータリングを行っている場面を抜き出したビデオを 4 本用意した。また、チューターとチューティの組み合わせが同じシーンも含まれている。

それぞれのチュータリングビデオは以下の観点からラベル付けを行う。

- チューターの発話
- チューティの発話
- チューターの作業空間への手の入り込み
- チューティの作業空間への手の入り込み
- チューターの注視方向（教材・相手の顔）
- チューティの注視方向（教材・相手の顔）

作成したラベルデータは、Sumi らが提案した iCorpusStudio[13] へ読み込む。そこでラベルデータ間の演算を利用し、次のラベルデータを作成する。

- チューターの沈黙（1000msec）を破った発話
- チューティの沈黙（1000msec）を破った発話
- チューティが発話と作業空間への手の入り込みを同時に行う

#### 4.5 各チュータリングビデオについて

ラベル付けを行った結果得られた各チュータリングビデオ内で見られる非言語情報の割合を図 3 に示す。また、それぞれのチュータリングビデオの特徴を以下に述べる。

Video1 のチュータリングは、チューターの発話が多いチュータリングである。また、チューティの作業時にチューティがあまり発話せず、チューターが多く発話していることが特徴的である。さらに、1 回のチューティの発話時間

もチューターに比べて短い傾向が見られた。これらのことから、チューターに促されてチューティが作業をしているチュータリングという印象を受ける。

Video2のチュータリングは、チューティの発話と作業が多いチュータリングである。また、チューティの作業時に発話も同時に行っている場面が多く見られた。一方で、チューターは説明する際は発話と作業をしっかり行うが、チューティが発話や作業を行う際には発話量や作業量がほとんど見られなかった。これらのことから、チューティ主導でチューターはそのサポート役となったチュータリングという印象を受ける。

Video3のチュータリングは、前半と後半で発話と作業の主体が入れ替わるチュータリングである。前半はチューターの、後半はチューティの発話と作業が目立った。また、チューティの作業と発話は同期していた。これらのことから、チュータリングの中盤でチューターからチューティへ大きく主導権が渡されるチュータリングという印象を受ける。

Video4のチュータリングは、終始チューターの発話と作業が見られるチュータリングである。チューティの発話は非常に少なく、作業はほとんど行っていない。これらのことから、チューターからチューティに対して知識を伝達するチュータリングという印象を受ける。

各チュータリングビデオで見られた非言語情報の割合

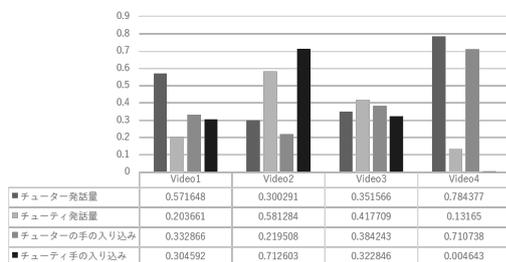


図 3 各チュータリングビデオで見られた非言語情報の割合

## 5. アンケート調査結果

### 5.1 チュータリングビデオに対する印象評価

今回行ったアンケート調査で、17件の回答が得られた。内訳は、メタ学習ラボに所属しているチューターから10件、一般の学生から7件であった。はじめに、メタ学習ラボ所属チューターと一般学生の2群に分け、それぞれのチュータリングビデオに対する評価を図4に示す。

図4の凡例は左から、チューティに会話の支配権があったか・他人に勧められるチュータリングであったか・チューティに自発性があったか、を表している。図4から、メタ学習ラボ所属チューター(図4内、MLLラベル)の評価値の分散が思いのほか大きいことがわかる。学習支援組織に所属せず、チュータリングに関する訓練を受けていない

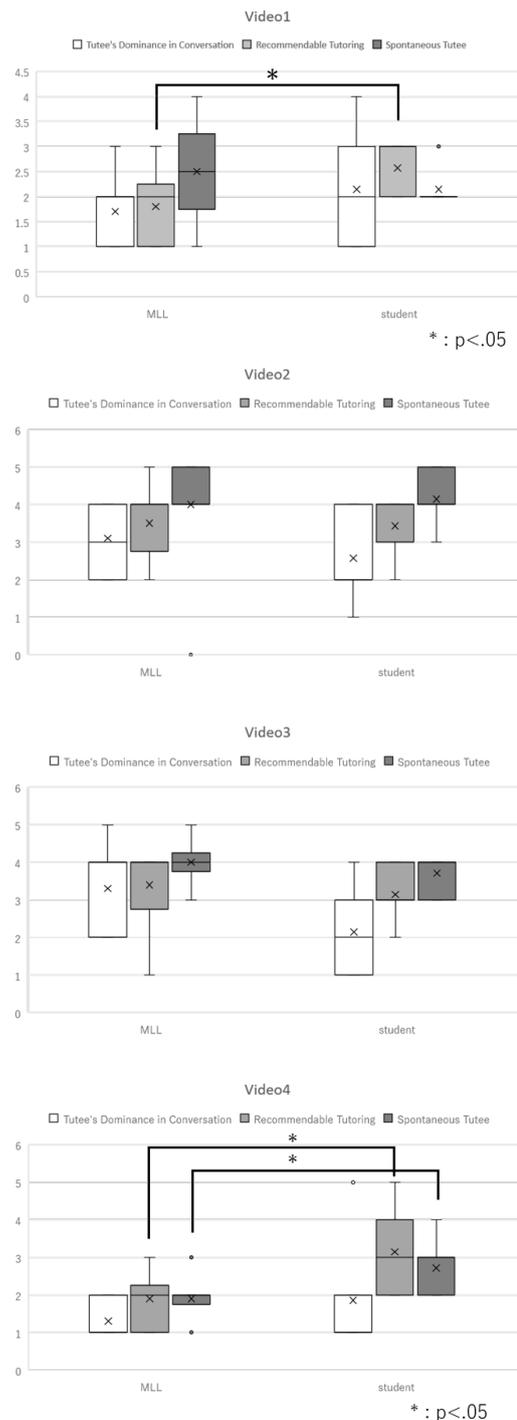


図 4 各チュータリングビデオに対して3つの観点からの評価値を示した箱ひげ図

一般学生の分散が大きくなることは容易に想像できる。しかし、メタ学習ラボ所属チューターの評価値においても分散が大きくなることは、筆者の想像と異なる。このような結果となった原因として、そもそも見てもらうチュータリングビデオのサンプル数が少なかったことが考えられるため、より多くのデータを収集し議論を深める必要があると考える。

次に、各チュータリングビデオの評価項目に関して、メ

タ学習ラボ所属チューターと一般学生の間で評価値が異なるか調べるために、分散が等しくないと仮定した2標本によるt検定を実施した。その結果、Video1における他人に勧められるチュータリングであったかという評価(A)・Video4における他人に勧められるチュータリングであったかという評価(B)・Video4におけるチューティに自発性があったかという評価(C)についていずれも有意差が得られた( $p < .05$ )。さらに、t検定で有意差が得られた群に対してマン・ホイットニーのU検定を行った結果、同様に有意差が得られた( $p < .05$ )。それぞれ得られたp値(危険率)については表1に示す。

表1 各検定で得られた結果  
t検定(両側検定) マン・ホイットニーのU検定

	t検定(両側検定)	マン・ホイットニーのU検定
A	p=0.0297	p=0.0476
B	p=0.0237	p=0.0181
C	p=0.0344	p=0.0275

次に、先ほど有意差が得られた評価項目を含むビデオについて具体的に議論を行う。各チュータリングビデオの内容や印象については4.5節を適宜参照してほしい。有意差が得られたチュータリングビデオの非言語情報について表2に示す。ただし、表2の単位はすべてチュータリング全体の時間で割ったものであり、表2の値は割合を示している。

表2 Video1と4で見られた非言語情報の割合

	Video1	Video4
チューターの発話量(sec)	0.572	0.784
チューティの発話量(sec)	0.207	0.132
チューターの手の入り込み(sec)	0.333	0.711
チューティの手の入り込み(sec)	0.305	0.005
チューティの発話と作業同期(sec)	0.056	0.005
チューターが沈黙を破った発話回数(回/sec)	0.038	0.041
チューティが沈黙を破った発話回数(回/sec)	0.027	0.008

表2から、Video1とVideo4どちらのチュータリングもチューターに発話量が偏っていることがわかる。また発話量が多いことから、チューターが沈黙を破った回数も多くなっている。加えて、他人に勧められるチュータリングであるかという観点からVideo1と4を評価した場合、メタ学習ラボ所属チューターが有意に低くスコアをつけている。従って、メタ学習ラボ所属チューターはチューターが多く発話を行ったり、沈黙を破って会話の主導権を握ったりすることに対してマイナスのスコアをつけていると予想できる。逆に、チューターが多く発話してチュータリング中の会話をリードした方が良いチュータリングになると一

般学生が思っているが、メタ学習ラボ所属チューターはそう思わないと言い換えることもできる。

他に、チューティに自発的に学ぼうとする姿勢が感じられたかという観点からVideo4を評価した場合、メタ学習ラボ所属チューターが有意に低くスコアをつけている。この時、チューターとチューティの作業空間への手の入り込みを見てみると、チューティの作業空間への手の入り込みが圧倒的に少ない。つまり、メタ学習ラボ所属チューターはチューティの自発性を評価する際にチューティの作業量に着目している可能性が示唆された。しかしながら、ビデオのサンプル数が少ないため、深い議論をするためにはより多くのチュータリングビデオに対して評価してもらう必要があると考える。

## 5.2 チュータリングを評価する際の判断材料

アンケート調査で得られたチュータリングを評価する際の判断材料において、チュータリングの進め方や発言内容などの言語的記述に比べて非言語的な記述が多く見られた。例えば、会話の支配権については、発話量だけでなく何ターン発話権がチューターからチューティに与えられたか着目するといった記述が見られた。また、自発的に学ぼうとする姿勢が感じられたかという観点から評価するために、チュータリングで利用する教材がチューターとチューティのどちらにあるか着目した記述が見られた。

一方で、他人に勧められるチュータリングであるかを評価する場合は、他の2つの観点に比べて言語的な判断材料が見られた。例えば、チューティに考えさせるチュータリングをしているかに着目した記述や、チューターがチューティに自ら学ぶように仕向ける発言や行動をしているかに着目した記述が見られた。ここに関しては、メタ学習ラボ所属の各チューターが持っている自分なりのチュータリングに対するポリシーや信念が関わってくるため内容に踏み込んだ記述が多く見受けられたと考える。ただ、今後は、チューティに考えさせるチュータリングといった非言語情報では表しがたいことに対して本研究のアプローチで説明していきたいと考える。

## 5.3 チュータリングスコアモデルの作成

本研究で得られたアンケート調査の結果と各チュータリングビデオにラベル付けされた非言語情報から、チュータリングスコアモデルの作成を試みた。目的変数をアンケート調査で得られたリッカート尺度の値、説明変数としてラベル付けされた表2で示したモダリティを利用し重回帰分析を行った。しかしながら、全観測数は40あったものの、説明変数の数に対してチュータリングビデオのサンプル数が4つと非常に少ないため、うまく線形式を得られなかった。今後、説明変数の数の10倍のサンプル数を確保し、チュータリングモデルの作成や作成されたモデルの評価を

メタ学習ラボ所属チューターに行ってもらおう予定である。

## 6. おわりに

本稿では、チュータリングスコア算出モデルを作成するための取り組みの一つとして、あるシーンに分割されたチュータリングビデオをチュータリングを専門の学習支援組織に所属するチューターにスコア付けしてもらい、ビデオ内に見られた非言語情報との関連性やチューターが持つ評価の判断基準について議論した。アンケート調査の結果、メタ学習ラボ所属チューターと一般学生が持つチュータリングに対する評価軸に差があることが示唆された。さらに、メタ学習ラボ所属チューターがチュータリングを評価する際に重視している観点をアンケート調査とチュータリング内に見られた非言語情報から分析した結果、発話量や作業量を重視していることが示唆された。

非言語情報がチュータリングの評価に与える影響についてさらに詳しく調査するために重回帰分析を試みたが、良い結果が得られなかったため今後さらにチュータリングビデオの数を増やす必要がある。また、チュータリング内の非言語情報を自動的にラベリングするシステムも今後実装していく必要があると考える。

## 参考文献

- [1] 福島悠介, 山崎俊彦, 相澤清晴. 文書と音声解析に基づくプレゼンテーション動画の印象予測. 電子情報通信学会論文誌 D, Vol. 99, No. 8, pp. 699–708, 8 2016.
- [2] 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.
- [3] 中園篤典, 谷川裕稔. アクティブラーニング批判的入門. 大学における学習支援への挑戦, No. 4. ナカニシヤ出版, 2018.
- [4] A. Kendon. *Gesture: Visible Action as Utterance*. Cambridge University Press, 2004.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.