

# 身体および生理データを用いた発言伝達強度の弁別モデル構築

竹平幸矢<sup>†1</sup> 遠田敦<sup>†2</sup>

**概要:** 人間が社会生活を営む中で重要な能力の1つとして挙げられるのはコミュニケーション能力である。しかし、人間同士の直接対話におけるコミュニケーションでは発言の受け取り方が人によって異なるため、意思伝達が正常におこなわれない可能性がある。そこで本研究では発言伝達強度の違いを示す弁別モデルの構築をおこなった。その結果、利き手の加速度、心拍数、集中値、リラックス値を用いた弁別モデルの予測精度が96.35%となり、他のモデルよりも高い値を示していた。そのため、この4種類を用いた弁別モデルが有用であるということが示された。

**キーワード:** コミュニケーション, 機械学習, RNN, 発言伝達強度

## Study on Discrimination Model of Speech Transmission Intensity using Physical and Physiological Data

YUKIYA TAKEHIRA<sup>†1</sup> ATSUSHI ENTA<sup>†2</sup>

**Abstract:** Communication is important ability of human in a member of society. But the communication has problem that interpretation way of speech is individual differences. Consequently, a lack of understanding between the parties will cause. This research is study on discrimination model between strongly speaking state and usual speaking state. This model used RNN of machine learning with physical and physiological data. As a result, the prediction accuracy of the discrimination model using dominant hand acceleration, heart rate, concentration value, relaxation value was 96.35%. The result of this model was higher than other models. Therefore, discrimination model using the 4 types showed that useful.

**Keywords:** Communication, Machine Learning, RNN, Speech Transmission Intensity

### 1. はじめに

人間が社会生活を営む中で、重要な能力の1つとして挙げられるのはコミュニケーション能力である。互いの意思を伝達し合うことで手段、行動の多様化を行い、利害関係の一致、目的の達成を行うことができるためである。人間の文明の発展に伴って、異なる風土や文化的背景を持つ人々が互いにコミュニケーションをとるため、様々な身体の進化や言語、文字が発達してきた。

現代においてもコミュニケーション能力は重要視されている。例えば企業の採用選考時では、図1のように、コミュニケーション能力が最も重要とされていること[1]や、文部科学省がコミュニケーション能力を向上させる政策を実施していること[2]が挙げられる。

コミュニケーションに関する研究は数多く存在し、機械を用いたコミュニケーション能力の支援、向上を目的とする研究も盛んになされてきた。例えば機械を介して遠隔地の人間とのコミュニケーションをとろうとするもの[2]、人間と機械とのコミュニケーションをとろうとするもの[3][4]などが挙げられる。また人間同士のコミュニケーション支援を目的とした会話ロボットの研究もなされている。藤江らは被験者に対して会話を中心としたゲームをおこな

わせた[5]。この時、会話ロボットを被験者と同じ立場の1人として参加させ、この会話ロボットがどれだけコミュニケーションを支援できるかを検証した。結果によると、ロボットのおこなった回答や、独り言によって、被験者の笑顔時間増加や、アンケートの回答から楽しさが増加したことが示された。これによりコミュニケーションの活性化が確認されたと報告されている。

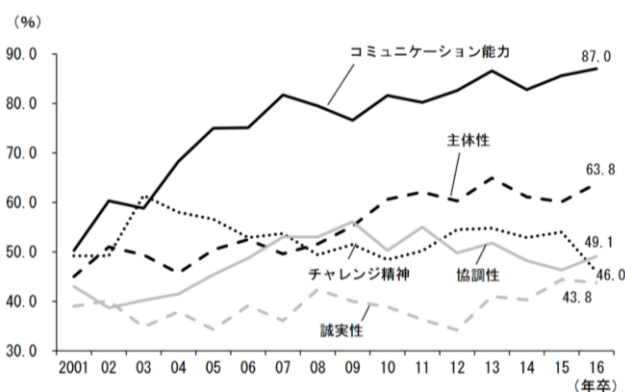


図1 企業採用選考時に重視する要素(上位5項目)の推移  
Figure 1 Changes in Elements to Emphasize (Top 5 Items) when Selecting Companies for Recruitment.

<sup>†1</sup> 日本大学大学院生産工学研究科  
Graduate school of Industrial Technology, Nihon University  
<sup>†2</sup> 日本大学  
Nihon University

また会話から状態を推量する研究もなされている。M. Swertsらは、自分の質問に対して相手が答えられるかどうかを推量する能力 (Feeling of Another's Knowing) に関する研究をおこなった[6]。これにより、表情などの視覚情報と、イントネーションなどの音声情報によって推量できることが示された。さらに Sinclar.Jらによって、言語や表情の分析によって、会話の状態を推定する研究もなされている[7]。

しかし一方で、人間同士のみの直接対話状況におけるコミュニケーションに関して、言語ではなく身体や生理データによって機械が人間の発言状態を推量できるのかという研究は十分になされていない。

人間同士の直接対話における、コミュニケーションの問題点として、発言の受け取り方が人によって異なることが挙げられる。これは自分が相手に対して強く、重くまたは弱く、軽く発言したにも関わらず、相手にはその逆、または別の発言が重く受け止められてしまう現象のことを指している。この現象により、お互いに重視する発言が異なるため、意思伝達が正常におこなわれない可能性がある。そこで本研究では、人間同士の直接対話において、発言伝達強度の違いを示す、弁別モデルの構築を目的とした。弁別モデルの構築には、発言者の身体および生理データを用いることとした。またそのデータを取得する器具を開発した。

## 2. 研究目的

本研究では、身体および生理データを取得する器具を開発し、得られたデータを用いて、発言伝達強度の違いを示す、弁別モデルの構築を目的とする。

## 3. 研究方法

### 3.1 実験概要

本研究の実施に先立ち、装着者から身体および生理データを取得する器具を開発した。被験者にはこの器具を装着させ、日常会話程度の伝達の強さ (通常伝達) と相手に対して最も伝えなければならない程度の伝達の強さ (強化伝達) をおこなわせた。この時、実験者からは発言の内容が書かれた原稿と、伝達の見本が提示され、それらを参考に発言をおこなわせた。被験者の発言中の身体および生理データを記録し、分析をおこなった。取得するデータとしては、利き手の加速度、心拍数、脳波による集中度とリラックス値、声の音圧の5つとした。発言を行う時間に制限は設けず、1人あたり約30分間の実験を32名に対して実施した。

### 3.2 実験場所

本実験は、大学キャンパス内の一室を用いて実験を実施した。実験の様子を図2に示す。



図2 実験の様子

Figure 2 The Conditions of an Experiment.

### 3.3 被験者

デザイン系学科に在学する大学生 (20~22 歳)、性別は問わず、32名に対して実施した。

### 3.4 実験器具

本実験に用いる実験器具は2種類に分けられている。まず1種類目として利き手の加速度、心拍数、脳波による集中度とリラックス値の取得に用いた器具である。これは ESP-WROOM-32, Arduino, microSD カードモジュール、心拍センサ、脳波センサ、加速度センサを接続した実験器具である (図3)。脳波センサに関しては、NeuroSky 社の物を使用した。簡易的ではあるが、日常生活において装着可能であることを理由に選定した。

2種類目の器具は声の音圧の取得に用いた器具である。これは Microsoft 社の Surface pro3 を主体とし、Sound Blaster X-Fi Go! Pro r2 (16bit, 44.1kHz, モノラル) と付属のマイクを用いて声の音圧の録音をおこなった。その実験器具を図4に示す。



図3 心拍、脳波、加速度器具

Figure 3 Heart Rate, Electroencephalogram, Acceleration Device.



図 4 録音器具

Figure 4 Recording Device.

### 3.5 計測方法

本実験に用いた器具の計測方法をそれぞれ説明する。心拍数、脳波、加速度に関しては、それぞれのセンサから取得したデータ、経過時間をテキストデータとして microSD へと保存した。この時、心拍センサは被験者の耳たぶ、脳波センサは頭部、加速度センサは利き手の中指に装着させた。またスイッチを押すたびに、押したタイミングと、その回数をそれぞれの microSD カードに保存した。このタイミングと回数、経過時間を用いて、データの照合をおこなった。声の音圧はマイクから得られた音圧を録音ソフトの SoundEngine Free を用いて録音した。この時、マイクに付属のクリップを用いて、被験者の首元にマイクを固定した。

### 3.6 教示内容

本実験では被験者に対して、架空の状況の一員であることを説明し、その状況に沿って指定された発言をおこなわせた。この時、発言させる原稿の内容としては、ある作品において、紹介する点（伝達事項）が 3 点用意された文章である。その 3 点の内、1 点のみ下線が引かれており、最も強く相手に伝達しなければならない部分とした。この最も強く伝達する発言を強化伝達発言、他の 2 点を通常伝達発言とした。また最も強く紹介すべき点が異なる原稿を 3 種類用意した。

実験者が発言の見本を提示した後、被験者に同様の発言をおこなわせ、その時の身体および生理データを記録した。実験者が提示する発言の見本では、強化伝達時に手を動かし、声が大きく、はっきり、ゆっくり、とした喋り方[8]を意識した。

原稿 1 種類ごとに 3 回発言し、3 種類すべてにおいて同様に実施した。分析に用いるデータは、3 回目に発言した時のデータとした。これは被験者の照れや不慣れ等をできる限りなくし、被験者が自然に強化伝達をおこなっている状態に近づかせるためである。また順序効果による慣れを考慮して、1 パターンごとに 5 人ずつ実験を実施した。被験者に提示する架空の状況、発言原稿文を以下に記す。

#### (1) 架空の状況について

あなたはある作品の制作者の一人で、その作品をあるコンテストへと提出しようとしている。コンテストは 3 次選考まで存在し、1 次では書類選考、2 次はプレゼンテーション選考、3 次は実物を用いた最終選考となっている。既に 1 次選考を通過し、あなたは 2 次選考にてプレゼンテーションをおこなうこととなっている。今からその練習を始めるところであり、あなたは作品の発表者、実験者は審査員の役である。作品介绍原稿上の下線部分が制作した作品に関して最も伝えたい部分である。

#### (2) 発言原稿文について

「今回、私のチームが開発、制作した作品は掃除機です。」「この掃除機は台風のようなサイクロン構造を用いており、強い吸引力が長く続きゴミを吸い続けます。（伝達事項①）」「形状は人間工学に基づいて設計されているため、使いやすく、部屋の隅々まで掃除機をかけることができます。（伝達事項②）」「製作では実際に動き、実用できるレベルにまですることが出来たため、次の選考に進んだ際には作品の性能や使いやすさを体感することができます。（伝達事項③）」

この時の実験変数を表 1 に、原稿と伝達事項の組み合わせを表 2 に示す。

表 1 実験変数

Table 1 Experimental Variables.

総伝達事項数	強化伝達事項数	パターン数
3	1	6

表 2 伝達事項の組み合わせ

Table 2 Combination of Transmission Matters.

原稿	伝達事項①	伝達事項②	伝達事項③
A	強化	通常	通常
B	通常	強化	通常
C	通常	通常	強化

### 3.7 実験手順

今回の実験でおこなった手順を以下に示す。

- ① 被験者と実験者を、机を挟んだ状態にて着席させ、発言時における身体および生理データを取得する実験という教示をおこなった。
- ② 被験者の名前、年齢、性別、利き手を聞き、記録した。
- ③ 被験者に対して、心拍、脳波、加速度器具、録音器具、を実験者の指示に従って、装着させた。それと同時に、実験者の PC に表示させている音圧の録音と、温度の録画を開始した。

- ④ 被験者に対して、架空の状況の一員であることを教示し、1種類目の原稿を提示した。
- ⑤ 実験者が見本の発言を1回おこない、続けて被験者に発言を1回おこなわせた。
- ⑥ ⑤を同様に2回おこない、1種類目の原稿を計3回発言させた。
- ⑦ 2種類目の原稿を提示し、⑤、⑥と同様の手順をおこなった。
- ⑧ 3種類目の原稿を提示し、⑤、⑥と同様の手順をおこない、実験を終了とした

### 3.8 分析方法

本実験の分析では、機械学習の手法であるニューラルネットワークを用いて、弁別モデルの構築をおこなった。まず表計算ソフトの Excel を用いてデータの整理を行い、そのデータを用いて Python のライブラリである TensorFlow を使用した。

取得した身体および生理データの要素において、用いた値についてそれぞれ記述する。利き手の加速度は  $x$ ,  $y$ ,  $z$  の三軸から得られる3方向のベクトルを、1つに合成した値を用いた。心拍数では、1分間の拍動の数を示す BPM の値が得られるため、その値を用いた。脳波による集中値とリラックス値は、実験器具から得られた値を用いた。この2つの値は最小値0、最大値100で与えられるものである。弁別モデルの構築は、5種類の身体および生理データのうち、1-5つを選んだ、すべての組み合わせ、計31パターンをそれぞれの被験者に対して、おこなった。この5種類のデータにおいて、取得時間はすべて同一であるが、データ数はそれぞれ異なるため、弁別モデルの構築の際には、線形補間によって、最もデータ数の多いものにデータ数を合わせた。

得られたデータのうち9割をモデル構築に用いる教師データとし、残りの1割は、構築されたモデルの弁別の予測精度を算出するために用いるテストデータとした。その結果から、どの種類の組み合わせを用いた弁別モデルが最も高い予測精度を示すのか、またその予測精度はどれほどの値を示すのかを明らかにした。

## 4. 実験結果

実験により、26人分の身体および生理データを取得し、データ整理およびモデルの構築をおこなった。残りの被験者6人の実験では、今回開発した身体および生理データを取得する器具が正常にデータを取得することができなかったため、モデルの構築をおこなわなかった。

本実験より得た、被験者ごとに算出した予測精度を表3に示す。用いた身体および生理データは、利き手の加速度は加、心拍数は心、集中値は集、リラックス値はリ、声の音圧は音と表記した。この表は最も濃い緑色はすべての値と比べて最も低い値、最も濃い赤色はすべての値と比べて

最も高い値を示している。考察を以下の3つの項目ごとに記述する。

### (1) 身体および生理データの要素数について

予測精度は用いた身体および生理データの要素数が多くなるにつれ、被験者全体として予測精度が高くなっていることがわかる。しかし、5つの要素すべてを用いた弁別モデルと、利き手の加速度、心拍数、集中値、リラックス値の4つの要素を用いた弁別モデルとでは、被験者26人中25人が、4つの要素を用いた弁別モデルの予測精度の方が5つの要素すべてを用いた弁別モデルの予測精度よりも高い。このことから、単純に弁別モデルの構築に用いる身体および生理データの要素数を多くすれば、より高い予測精度を算出できるということではないと考えられる。

### (2) 身体および生理データの組み合わせについて

それぞれの被験者の最も高い予測精度を示した表4より、26人中、10人が利き手の加速度、心拍数、集中値、リラックス値を用いた弁別モデルが最も高い予測精度である97[%]前後を示した。また、これに該当しなかった他の被験者16人のうち、15人においては、それぞれの最も高い予測精度の値と、利き手の加速度、心拍数、集中値、リラックス値を用いた弁別モデルの値との差は5[%]前後であった。

次にこの利き手の加速度、心拍数、集中値、リラックス値を用いた弁別モデルに関して着目する。予測精度の平均値を示した図5より、このモデルの予測精度の平均値は最も高い96.35[%]を示していることがわかる。予測精度の標準偏差を示した図6においては、3番目に低い値を示しているため、他のモデルよりも値のばらつきが低く、安定して高い予測精度を示すことがわかる。

これらのことから、最も高い予測精度を持つ弁別モデルは、個人ごとに異なるが、利き手の加速度、心拍数、集中値、リラックス値を用いた弁別モデルは、いずれの被験者においても高い予測精度を示していることがいえる。それにより、予測精度の高い、発言伝達強度の弁別モデルを構築するには、利き手の加速度、心拍数、集中値、リラックス値を用いることが有用であると考えられる。

### (3) 予測精度の向上について

弁別モデルの予測精度にて、最も低い値は約33.33[%]、最も高い値で100[%]を示した。100[%]である弁別モデルは、心拍数、集中値、リラックス値を組み合わせたものであり、これらは他と比べ、取得データ数が少なく、弁別モデルの予測精度に用いるテストデータは少ないことから、100[%]という値を算出したと考えられる。そのため、取得データ数を多くすることによって、より正確な予測精度が算出できると考えられる。

また、通常伝達を0、強化伝達を1とし、2種類の値しか用いなかった。これにより0または1以外の値を用いることによって通常伝達と強化伝達の弁別モデルの精度を上げることが出来ることが考えられる。

表 3 すべての弁別モデルの予測精度 (単位: %)

Table 3 Prediction Accuracy of All Discrimination Models. (Unit:%)

要素数	1					2					3					4					5									
	加	心	集	リ	音	加心	加集	加リ	加音	心集	心リ	心音	集リ	集音	リ音	加心集	加心リ	加心音	加集リ	加集音	加リ音	心集リ	心集音	心リ音	集リ音	加心集リ	加心集音	加心リ音	加集リ音	加集リ音
被験者1	64.66	53.33	66.67	66.67	65.17	75.00	66.09	71.84	70.91	60.00	46.67	79.40	77.78	69.20	67.86	95.11	83.62	79.59	88.51	65.62	73.03	86.67	88.12	89.31	86.66	97.70	83.22	82.98	84.12	94.84

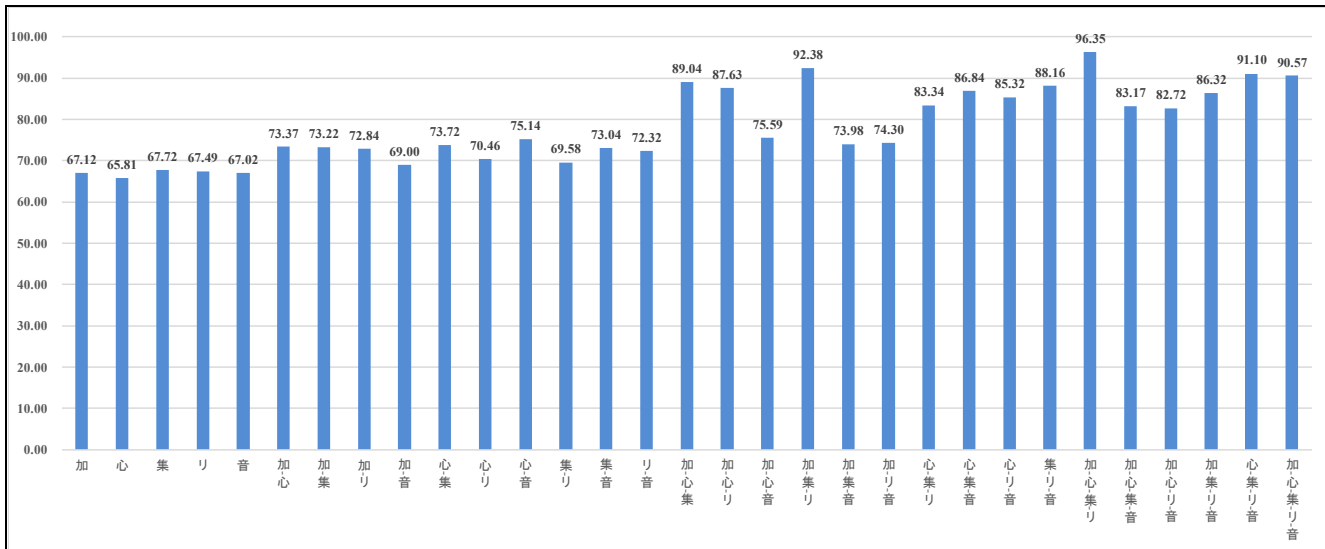


図 5 予測精度の平均値 (単位: %)

Figure 5 Average Value of Prediction Accuracy. (Unit:%)

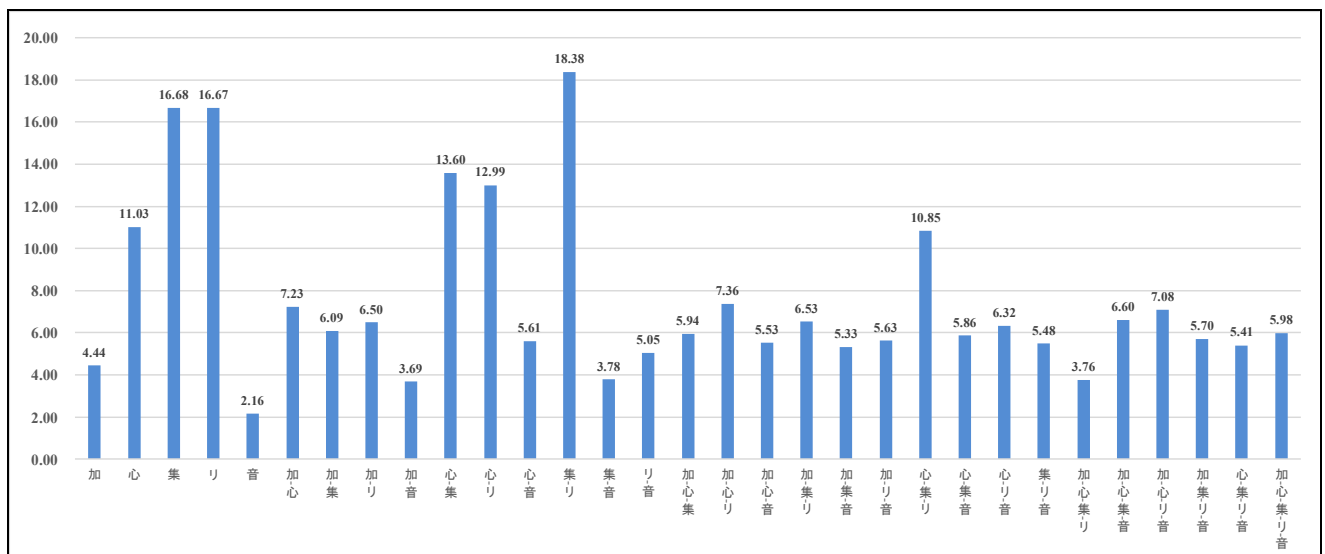


図 6 予測精度の標準偏差

Figure 6 Standard Deviation of Prediction Accuracy.

表 4 それぞれの被験者の最も高い予測精度 (単位: %)

Table 4 The Highest Prediction Accuracy of Each Subject.

(unit:%)

	最高予測精度	要素	加-心-集-リとの差
被験者3	97.70	加-心-集-リ	0.00
被験者4	92.33	加-心-集-リ	0.00
被験者5	100.00	リ	0.31
被験者6	99.10	加-心-集-リ	0.00
被験者7	99.56	加-集-リ	5.99
被験者8	99.48	加-集-リ	1.38
被験者9	100.00	集-リ	1.74
被験者10	100.00	心-リ	1.00
被験者12	100.00	心-集-リ	1.94
被験者13	99.70	加-心-集-リ	0.00
被験者14	97.87	加-心-集-リ	0.00
被験者15	100.00	集-リ	0.31
被験者16	99.71	加-心-集-リ	0.00
被験者17	100.00	心-集、心-集-リ	3.01
被験者20	94.22	加-心-集-リ	0.00
被験者21	93.45	加-心-集-リ	0.00
被験者22	100.00	集、集-リ	2.66
被験者23	99.15	加-集-リ	5.38
被験者24	95.91	心-集-リ-音	3.42
被験者26	100.00	心-集、心-集-リ	0.49
被験者27	100.00	心-集-リ	7.71
被験者28	99.47	加-集-リ	0.53
被験者29	92.35	加-心-集-リ	0.00
被験者30	100.00	心-集-リ	1.66
被験者31	99.00	加-心-集-リ	0.00
被験者32	99.71	加-集-リ	16.03

## 5. まとめ

本研究では、身体および生理データを取得する器具を開発し、得られたデータを用いて、発言伝達強度の違いを示す、弁別モデルの構築をおこなった。これにより明らかになったことを以下に示す。

- 被験者 26 人中 25 人において、身体および生理データの要素数が 5 つを用いた弁別モデルよりも、4 つを用いた弁別モデルの方が、予測精度が高い結果となった。そのため、単純に要素数を多くすれば予測精度が高くないと考えられる。
- 今回得た身体および生理的データでは、最も高い予測精度を持つ弁別モデルは、個人ごとに異なるが、利き手の加速度、心拍数、集中値、リラククス値を用いた弁別モデルは、いずれの被験者においても高い予測精度を示していた。そのため、利き手の加速度、心拍数、集中値、リラククス値を用いた弁別モデルが通常伝達と強化伝達を弁別するのに有用である。

今後、本研究で得た結果をもとに、発言伝達強度の弁別

モデルを構築する際、考慮すべき点を 4 つの項目に分けて以下に記す。

### (1) 発言時間の長期化

本実験にて得た身体および生理データは、被験者が発言していた時のものを使用したが、発言時間は、1 つの原稿につき 30 秒程度で、3 つの原稿を合わせても、90 秒程度の発言時間であった。そのため、今後は長期の発言を考慮する必要がある。

### (2) 身体および生理データの要素の追加

本研究においては、身体および生理データを利き手の加速度、心拍数、脳波による集中値とリラククス値、声の音圧の 5 つとして、弁別モデルの構築を実施した。この 5 つ以外のデータを用いることによって、予測精度が高い新たな組み合わせを明らかにすることができると考えられる。

### (3) 伝達状態の細分化

本研究のモデル構築の際には、通常伝達を 0、強化伝達を 1 として数値を与えた。しかし 2 種類だけでなく、複数の種類に細分化された伝達状態における、弁別モデルを構築することによって、より多様な状態を弁別することが可能になると推測される。

### (4) モデル構築の別手法

本研究のモデル構築には、Python に TensorFlow を用いた。この手法以外を本実験では実施できなかったため、別手法を実施し、本研究の手法と比較することにより、最適な弁別モデルの構築方法を明らかにすることができる。

## 参考文献

- [1] 日本経済団体連合会. 2016 年度 新卒採用に関するアンケート調査結果の概要. 2016-11-15, p. 1.
- [2] 柏原忠和, 大澤博隆, 篠沢一彦, 今井倫太. 遠隔コミュニケーションの為のウェアラブルロボットアバタの提案. 第 73 回全国大会講演論文集, 2011, p. 237-238.
- [3] 小野哲雄, 今井倫太, 石黒浩, 中津良平. 身体表現を用いた人とロボットの共創対話. 情報処理学会研究報告知能と複雑系 (ICS), 2001, p. 1-6.
- [4] 滝僚平, 前田陽一郎, 高橋泰岳. ロボットの情動行動が人間に与える影響の脳波解析. HAI シンポジウム 3A-2, 2010, p. 1-6.
- [5] 藤江真也, 松山洋一, 谷山輝, 小林哲則. 人同士のコミュニケーションに参加し活性化する会話ロボット. 電子情報通信学会, 2012, p. 37-45.
- [6] M. Swerts, E. Kraemer. Audiovisual prosody and feeling of knowing. Journal of Memory and Language, 2005, vol.53, p. 81-94.
- [7] Sinclair, J., M. Coulthard. Towards an Analysis of Discourse: The English Used by Teachers and Pupils. Oxford University Press, 1975.
- [8] 荒巻基文. 社会人のための伝える力. 産業能率大学出版部, 2013, p. 149-151.