

## 非タスク指向対話システムにおけるからかいユーモア応答の生成

Friendly-teasing Humor Response Generation  
for Non-task-oriented Dialogue System

上垣 貴嗣 (UEGAKI, Takatsugu) 指導：菊池 英明

## 1. はじめに

人との雑談を行う非タスク指向対話システム研究の目的の一つとして、友人同士のようなコミュニケーションの達成が挙げられる。この目標を達成するためには、システムがユーザーに対して距離感を詰める発話を行う必要がある。円滑なコミュニケーションについての理論であるポライトネス理論において、相手との距離感を詰める発話戦略はポジティブ・ポライトネス戦略と呼ばれる。ポジティブ・ポライトネス戦略には、からかいやいじり、皮肉といった攻撃性を持つユーモア（からかいユーモア）が含まれている。対話システムのからかいユーモア応答が、ユーザーに親しみを感じさせ、「またこのシステムを利用したい」という対話継続欲求を向上させることができると考える。そこで本研究では、からかいユーモア応答を生成するシステムを開発し、対話継続欲求への影響を調べる。

## 2. からかいユーモアコーパスの構築とデータ拡張

まず、対話文脈に対してからかいユーモア応答を付与したコーパス（からかいユーモアコーパス）を構築する。応答付与の対象とする対話データには、Twitterから取得したツイート・リプライ連鎖を利用する。プロの放送作家4名とお笑い芸人1名にツイート・リプライ連鎖に対するからかいユーモアの応答作成を依頼した。作成した応答と元データ合わせて3478件に対しクラウドソーシングにて「からかいユーモアらしさ」「暴言らしさ」の5段階評価（1-5）を行い、各得点の平均値を「からかいユーモアスコア」「暴言スコア」としてコーパスに付与した。

からかいユーモアコーパスを利用した対話応答生成モデル学習に向けて学習データを拡張するため、ツイート・リプライ連鎖に対してからかいユーモアスコアと暴言スコアを予測する回帰モデルを構築した。モデルにはRoBERTaを、学習データにはからかいユーモアコーパスを利用した。モデル精度のRMSEはからかいユーモアスコアで0.624、暴言スコアで0.547となり、ツイート・リプライ連鎖約292万件の内124,879件がからかいユーモアとして得られた。

## 3. 対話応答生成モデルの構築

図1に提案手法の流れを示す。提案手法は大きくエンティティ・リンク部と応答生成部に分けられる。

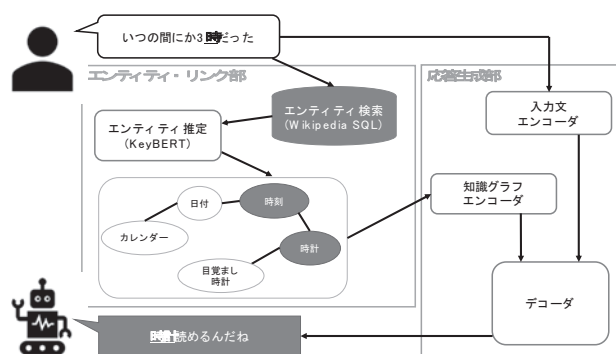


図1：提案手法の流れ

エンティティ・リンク部では入力文解析として、ユーザー入力文中の主要語について、Wikipediaページリンクを元に作成した知識グラフと結びつける。応答生成部では、入力文と知識グラフを入力として応答を生成する。応答生成モデルのベースにはT5を、学習データセットにはからかいユーモアコーパスに加え、前章で得られたからかいユーモア群のツイート・リプライ連鎖を利用した。

## 4. システム評価実験

システムの生成能力評価には、からかいユーモアコーパスに含まれるからかいユーモア群の一部をテストセットとして用いた。提案手法、既存手法のT5とテストセット応答について「応答の自然さ」「応答のからかいユーモアらしさ」「対話継続欲求」を比較した。実験はクラウドソーシングで行い、被験者は20名で5段階評価（1-5）とした。

対話継続欲求への影響評価実験には、リアルタイムな対話を実施した。ベースラインには一般的応答を生成するT5モデルを用いた。実験はクラウドソーシングで行い、被験者は20名とした。

## 5. 結果

## 1) 対話応答生成モデル評価

提案手法は、テストセット応答との比較では全項目で劣ったものの、からかいユーモア生成モデル間の比較では、全ての評価項目で既存手法を上回る結果となった。

## 3) 対話継続欲求評価

テストセットに対する評価とリアルタイム対話評価の両方で、からかいユーモア応答モデルの提案手法と、一般応答モデルとの比較において、対話継続欲求に対する有意にポジティブな影響は確認できなかった。

## フリースタイルバスケットボールバトルにおけるプレイヤーの待ち動作に対して審査員が抱く印象「かっこよさ」と「巧みさ」の検討

大野 俊尚 (OHNO, Toshihisa) 指導：三嶋 博之

フリースタイルバスケットボール (以下FB) とは音楽に合わせてダンスのようにバスケットボールを操るパフォーマンスのことである。文化の特徴としてストリートダンス同様の「バトル」があり、ボールを操る技術の巧みさだけでなく、「かっこよさ」も評価に含めて審査員の主観で勝敗が決まる。では、審査員はパフォーマンスの何からかっこよさを判断しているのだろうか。事前インタビューで得られた「立ち振る舞い」という観点から、バトルで自分のターンを待つ様子から検討する。

**方法:** 実験参加者はFBバトル審査員経験者である男性6名であった。実験参加者ごとにランダムな順序で実験刺激を提示し、印象の評価をさせた。実験後、なぜその回答をしたのかを尋ねた。

**実験刺激:** MOTION by RADiCALを用いてYouTube上のバトル動画からターン待ち中の動作データを抽出し、3Dモデルを作成した。動画はパフォーマンスの本編を始めたと考えられる最初の技を行った時点を終了時間とし、開始時間は終了時間の15秒程度前となるようトリミングした。そのため、実験刺激には自分のターンを待つ「待ち」の様子と、自分のターンになり、ステージ中央に受かって歩いてからパフォーマンスを始める「入り」の様子が含まれていた。

**結果:** 実験参加者の評価において実験参加者間の信頼性を検定するため級内相関係を求めた結果、0.273(95%CI = 0.103-0.474)であった。実験参加者の評価は実験参加者間で一致しておらず、実験参加者ごとに異なっていた。

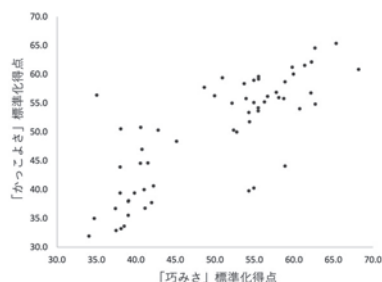


図1. 「かっこよさ」標準化得点と「巧みさ」標準化得点の散布図

実験参加者ごとに「かっこよさ」得点と「巧みさ」得点を標準化し、散布図を作成した。「かっこよさ」標準化得点と「巧みさ」標準化得点との相関係数は0.799 (95%CI = 0.627-0.884)であり、強い相関が見られた。しかし、決定係数は0.639であり、「巧みさ」が「かっこよさ」を、また

は「かっこよさ」が「巧みさ」を説明しきれないわけではなかった。「かっこよさ」を評価する観点と「巧みさ」を評価する観点とで差異があるためだと考えられる。

この評価の差異を検討するため、フリーのテキストマイニングソフトKH Coder 3 (樋口, 2021) を用いて印象評価実験後のインタビュー内容を検討した。

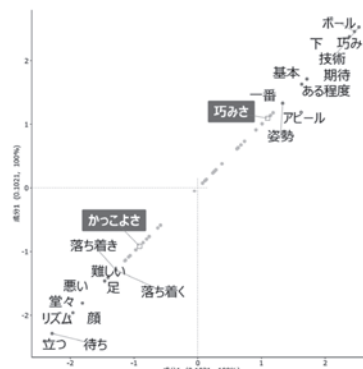


図2. KH Coderによる対応分析の出力結果 (原点から離れた上位20語のみラベル表示)

対応分析で抽出された語をKWICコンコーダンスで文脈の確認を行なったところ、「かっこよさ」は「待ち」の様子に関する語、「巧みさ」は「入り」の様子、ボールを扱う様子、体の使い方に関する語が特徴語であった。

**考察と展望:** 「かっこよさ」はターン待ち中であっても観客に見られている意識があるように感じる動作かどうかで判断されたのだと考えられる。また「巧みさ」に体の使い方に関する語が出たのは、ボールを扱う技術のみを評価しないというFBの文化的特徴が現れたと考えられる。

しかし、本研究では待ち中の1名しか表示していなかった。実際のバトル場面では清水ら (2018) のような対戦相手との協調運動も見られると考えられる。相手パフォーマンスとの協調関係の有無による印象の変化も検討する必要がある。

### 引用文献:

- 樋口 耕一 (2021). KH Coder 3 Retrieved from <https://kncoder.net/> (2022年1月12日)
- 清水 大地・岡田 猛 (2018). 舞台表現における共演者との相互作用—同期に着目した検討— 人工知能学会全国大会 論文集, 32, 1-4. [https://doi.org/10.11517/pjsai.JSAI2018.0\\_4B1OS19c04](https://doi.org/10.11517/pjsai.JSAI2018.0_4B1OS19c04)

# 強化学習を用いたキャラクターらしさを持つ雑談応答の生成

## Generating Open-Domain Dialogue Responses with Character-likeness Using Reinforcement Learning

清水 健吾 (SHIMIZU, Kengo) 指導：菊池 英明

### 1. はじめに

近年、人と極めて自然な雑談対話ができる大規模なニューラル対話生成モデルが登場した。これらの対話生成モデルがユーザに「また対話したい」と思わせるための方法として、特定のキャラクターらしさをシステム応答に付与する方法が考えられる。本研究は、強化学習を用いてニューラル対話生成モデルをfine-tuningすることで、特定のキャラクターらしさを持つ雑談応答を生成するモデルを提案する。Fine-tuningの際に、特定のキャラクターらしさを持つそれぞれ独立した発話データのみを必要とし、対話形式のデータを必要としない点で既存研究と異なる。

### 2. 提案手法

提案手法は以下の3つの段階で構成される。

1. Twitter疑似対話データセットとキャラクターツイートデータセットを構築する。
2. Twitter疑似対話データセットを用いて大規模言語モデルを学習することで対話生成モデルを構築する。
3. 強化学習によって対話生成モデルをfine-tuningすることで発話に特定のキャラクターらしさを付与する。

#### 2.1 データセットの構築

対話生成モデルを学習するために2種類のデータセットを構築した。Twitter疑似対話データセットは、約292万件のツイートとリプライの連鎖を収集したものである。ツイート・リプライの連鎖は疑似的な対話とみなすことができる。キャラクターツイートデータセットは、ある特定のキャラクターを模したアカウントのツイートを約3.5万件収集したものである。

#### 2.2 対話生成モデルの構築

Twitter疑似対話データセットを用いて日本語学習済みGPT-2[1]を学習することで日本語版DialoGPT[2]を構築した。GPT-2とは大規模言語モデルの1つであり、DialoGPTとはGPT-2を対話生成タスクに適用したモデルである。構築したDialoGPTは本研究のベースとなる対話生成モデルであり、評価実験においてはベースラインシステムとして扱った。

#### 2.3 強化学習によるfine-tuning

まず、強化学習における報酬を算出する報酬モデルを構築した。キャラクターツイートデータセットを用いてRoBERTa[3]を2値判別タスクで学習することによって、対象のキャラクターらしさを算出する報酬モデルを構築した。次に、報酬モデ

表1 提案システムが生成した対話

User	今日寒いけど、アライグマさんは寒くない？
System	寒いのだー！暖かくして過ごすのだ！
User	アライさんはもふもふしていて暖かそう
System	もふもふなのだー！
User	私はアイスが好きで、今日もアイス食べちゃった。
System	アイス美味しいのだー！アライさんも食べるのだー！

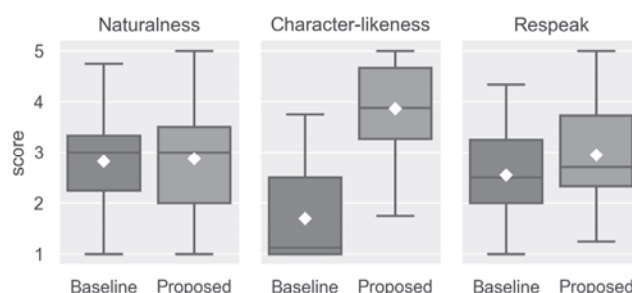


図1 評価実験の結果

(左から自然さ・キャラクターらしさ・対話継続欲求)

ルが算出するキャラクターらしさを報酬として、DialoGPTに対してProximal Policy Optimization[4]によって強化学習を行った。構築したモデルは評価実験において提案システムとして扱った。

提案システムが生成した対話の一部を表1に示す。

### 3. 評価実験

対象のキャラクターを知っている被験者50名が、ベースラインシステムおよび提案システムとチャットアプリ上で対話をした。対話に対して自然さ・キャラクターらしさ・対話継続欲求について5段階で評価した。結果を図1に示す。評価実験の結果、提案システムが、キャラクターらしさと対話継続欲求の項目において、有意に上回ることを確認した。

### 4. おわりに

本研究では、キャラクターらしさを持つそれぞれ独立した発話データを用いて強化学習を行うことによる、特定のキャラクターらしさを持つ対話応答を生成する手法を提案した。提案システムのユーザに、特定のキャラクターの印象を与え、「また対話したい」と思わせる効果を確認した。

### 参考文献

- [1] Radford et al. Technical report, Open AI, 2019.
- [2] Zhang et al. In *58th ACL*, pp.270-278, 2020.
- [3] Liu et al. *arXiv preprint*, 1710.03740, 2017.
- [4] Schulman et al. *arXiv preprint*, 1707.06347, 2017.

## 脳活動同期性に基づく道德教育の評価手法の開発

Development of Evaluation Method for Moral Education  
Based on Brain Synchrony

下田 香織 (SHIMODA, Kaori) 指導：松居 辰則

1. はじめに 現在の道德教育の効果は問題視されており、道德の授業の実施においては、学習者が教師の期待する正答を察知し、自身の感情を偽った意見を述べる傾向がある点など、様々な問題点が指摘されている[1][2]。指導方法の改善に至らない原因の一つとして、「効果的な道德の授業が実施されている状態」を客観的に評価する手法が存在しないことが挙げられる。一方、教師・学習者間で良質なコミュニケーションが行われた際、両者には高い調和感（授業がうまくいっているというポジティブな感覚）が生じると考えられる[3]。道德教育では、教師による一方的な情報伝達ではなく、教師・学習者間で、道德的価値観の「共有」が円滑に行われることが求められる。そこで本研究では、学習者の調和感に着目し、道德の授業の効果を生体計測に基づいて評価する手法の開発を試みた。

2. 道德の授業の効果と調和感の関連の検討 本研究では、道德の授業の効果と学習者の調和感の関連の検討を目的に道德の授業に取り組む学習者に対して質問紙調査含む実験を実施した。調査は、(1)大学生による道德の授業の再現、(2)高等学校で実施された道德の授業（学習者は高校生）、(3)大学で実施された道德の授業（学習者は大学生）の3つの形式で実施した。なお、本研究では道德の授業の効果を、「学習者が授業で取り扱われた道德的価値観について、授業を通してどれほど重要な内容だと感じられたか」と定義した。調査の結果、高等学校で実施された授業では、道德の授業の効果が高い学習者は、教師に対する調和感も高くなる傾向が見られた。このことから、「効果的な道德の授業が行われている状態」の評価にあたり、学習者の調和感評価を取り入れることが有用である可能性が示唆された。

3. 調和感と脳同期現象の関連について 前頭前野 (PFC) の活動は、他者の考えや感情を推し量ることと関連していると考えられており、教師・学習者間でのPFCの活動の同期は両者の良質な関係構築と関連する可能性が指摘されている[4]。そこで、授業中の学習者の調和感の高さと、教師・学習者間のPFCの活動の同期との関連を検討することを目的に、前節で述べた3つの形式の実験において、左右2チャンネル型のNIRS (HOT-1000, 株式会社NeU製) による測定を実施した。なお、本研究では脳活動の同期を、ウェーブレットコヒーレンスに基づいて評価した[4]。その結果、全ての形

式の実験において、学習者の調和感の高さによって、相互学習場面（教師・学習者が意見交換を行う場面）の左チャンネルの低周波帯の同期の程度が異なることが明らかになった。このことから、学習者の調和感は、教師・学習者間の脳同期現象に基づいて評価できる可能性が示唆された。

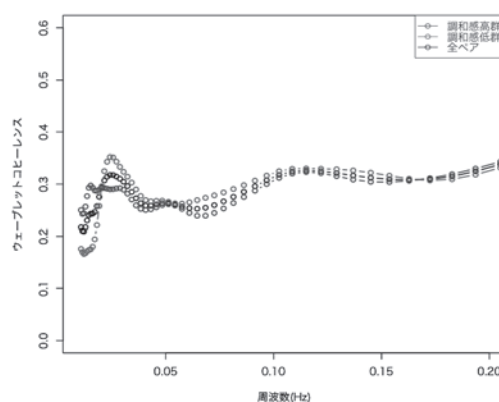


図1 相互学習場面の左チャンネルの脳同期  
(大学生による授業再現実験の場合)

4. まとめと今後の課題 本研究の結果、道德の授業の効果は、学習者に生じる調和感と関連し、教師・学習者間の脳同期現象によって評価できる可能性が示唆された。今後は、小・中・高等学校での知見を蓄積し、脳同期現象に基づく道德教育の評価手法の妥当性を検証することで、より良質な道德教育の提供に貢献することが期待される。

## 参考文献

- [1] 松尾直博: 道德性と道德教育に関する心理学的研究の展望-新しい時代の道德教育に向けて-, 教育心理学年報, Vol.55, pp.165-182 (2016)
- [2] 徳永正直, 宮嶋秀光, 榊原志保, 堤正史, 林泰成: 道德教育論: 対話による対話への教育, ナカニシヤ出版(2003)
- [3] Vanessa Rodriguez: The Human Nervous System: A Framework for Teaching and the Teaching Brain, *Mind, Brain, and Education*, Vol.7, No.1, pp.2-12 (2013)
- [4] Takayuki Nozawa et al.: Prior physical synchrony enhances rapport and inter-brain synchronization during subsequent educational communication, *Scientific Reports*, Vol.9, No.12747 (2019)

## 人の感性を考慮した自動歌唱評価システムの開発

## Automatic Singing Evaluation System Considering Human Sensitivity

高田 兼輔 (TAKADA, Kensuke) 指導：菊池 英明

## 1. 背景・目的

近年、カラオケや動画共有サイト、SNS 等での「歌ってみた」の普及により、人に歌唱を聴かせる機会が増えている。このような機会の増加に伴い、聴取者に「うまい」と思わせたい歌唱者にとって、歌唱評価のニーズが高まっている。

本研究では、[1] で明らかにした人が歌唱を聞いた際の印象評価モデルを基に歌唱技術のみにとらわれない、人の感性を考慮した自動歌唱評価システムの開発を目的とする。

## 2. 印象評価モデル

[1] で行なった共分散構造分析ではモデルとしての精度が低かったため、精度の向上を図った。①評価者傾向のばらつき抑制のためのデータ正規化、②クラスタリングによる評価者の限定、③主成分分析、因子分析を経た評価語の削減、の3段階を行い、新しくモデルを作成した。

## 3. 評価システムの開発

上記のモデルを基に自動歌唱評価システムの開発を行なった。本システムの流れは以下の通りである。

- ① 歌唱入力
- ② 入力された歌唱から音響特徴量 (20種類) を抽出
- ③ モデルを基に評価点 (100点満点) を計算
- ④ 評価点、高評価因子、低評価因子、アドバイスを表示

## 4. システム評価実験

本実験は、主にシステムの有用性を測るものだが、最も重要な点は「歌唱者が歌唱力の向上に足りていないものを把握できることが有意義かどうか」である。被験者を実験群と統制群に分け、実験群には感性評価を行う感性版 (図1) を用いて実験を行い、統制群には技術のみを評価する技術版 (図2) を用いて実験を行った。

実験は実験群・統制群共に男女5名ずつの計20名を対象に行い、被験者に歌唱評価→歌唱練習→歌唱評価→アンケートの流れで進めてもらった。

歌唱評価では被験者にそれぞれのシステムを用いて評価してもらった。歌唱練習ではそれぞれのシステムを用いて歌唱の練習を行ってもらった。アンケートでは、システム

の評価の正当性、システムの有用性、使用継続欲求などに加え、自己効力感尺度について聞いた。



図1 感性版の表示例

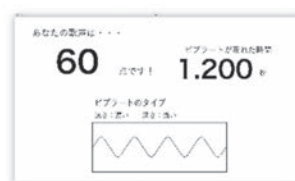


図2 技術版の表示例

## 5. 実験結果

アンケートの回答をt検定を用いて分析したところ、どの質問にも有意差が見られなかった。よって、システムを用いる効果、短時間で歌唱力上達、使用継続欲求には技術版と感性版の間に差がないという結果になった。しかし、システムとしてのうまさ評価の正統性に関しても、技術版と感性版の間の評価に差がなかったと考えられるため、ある程度の正統性があることが明らかになった。

次に自己効力感尺度について分析を行った。両システムの違いによる効果が統計的に有意か確かめるために、有意水準5%で両側検定のt検定を行った結果、統制群は $t(3) = 0.867$ で有意差なし、実験群は $t(4) = 2.986$ でサンプル数は少ないながらも有意差ありと判断できる。統制群と実験群はともに実験前よりも実験後の方が自己効力感評価はあがっているため、実験群は感性版システムを用いることで自己効力感が向上すると推測ができる。

## 6. まとめ

「歌唱者が自信の歌唱に足りていないものを直感的に把握することができる」ことの有用性、「感性版システムは技術版システムに比べて、短時間で歌唱力が上がり、使用継続欲求が高まる」ことについての検討を以下にまとめる。

- ① 感性版と技術版のシステムとしての評価には有意な差が見られなかった。
- ② 感性版システムは技術版システムに比べて自己効力感向上効果が高い。

## 参考文献

- [1] 高田兼輔ら, "歌唱のうまさに対する印象評価構造", 情報処理学会研究報告音楽情報科学, 2019

# タッチパネル携帯端末において指を離す際の動作が ポインティング特性に及ぼす影響

## An Effect of Land up Action in Pointing Feature with Touch-Sensitive Screens

陳 揚 (CHIN, You) 指導: 藤本 浩志

### 1. 背景及び目的

近年投影型静電容量方式タッチパネルの普及率が向上し、タッチパネルはスマホだけではなく、多くの製品に応用されている。しかしながら、ドラッグ操作において指先をモニターから離す際にスライダサムの意図しない移動が起こる場合がある。より精細な操作を実現するためには、この意図しない移動が発生する現象の原因を研究しなければならない。指先は粘弾性特性を持つ組織である。つまり、指先が画面に接触するとき、皮膚及び皮下組織の変形が起こる。この変形が、ズレを引き起こる原因かもしれないと考えられる。そこでタッチパネル携帯端末におけるドラッグ操作性向上を指向し、上述したドラッグ操作末期、ズレが発生する課題に対して、本研究では人の指先がタッチパネル端末から離れる際の動作や姿勢によって、指先の接触面及びポインティングの中心位置の変化傾向がズレの発生にどのような影響があるかを明らかにすることを目的とした。

### 2. 実験

実験参加者は、大学生17名（平均年齢 $24 \pm 2.3$ 歳）であった。実験装置として、以下の図1の装置を利用した。

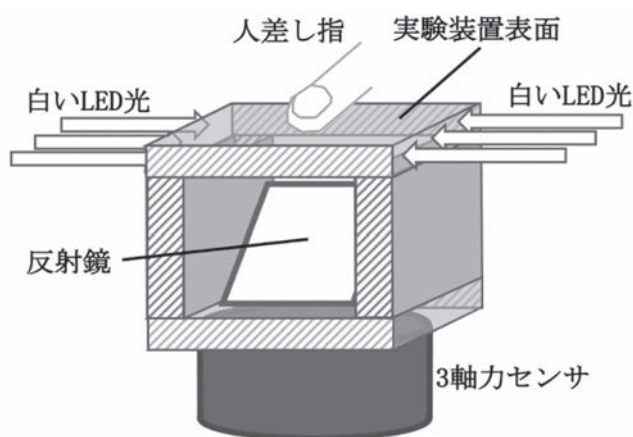


図1 実験装置

また本実験では、ドラッグ操作末期に指先がタッチパネルから離す際の姿勢の違いにも着目するので、5種類の姿勢の違いを実験因子として採用して実験を行った。

### 3. 結果及び考察

結果を図2に示す。姿勢3, 4と5については平均ズレ量が大きかった。他方、姿勢1と姿勢2の全体的ズレ量は小さい傾向が見られた。特に姿勢1はドラッグ操作の逆方向に移動する傾向があつて、姿勢2によりズレが高い傾向が見られた。姿勢2は平均ズレが少なく、ほぼズレでいないことが確認出来た。

### 4. 結論

右手の差し指を利用して左から右へドラッグ操作を行う場合、姿勢1、姿勢2の場合は他の姿勢より、指先を離れ始めるから完全に離れるまでズレが小さいという結果が分かった。更に姿勢2は姿勢1より、顕著的にズレを抑える傾向が明らかにした。図2から、5種類の姿勢ごとのズレの特徴が明らかになった。

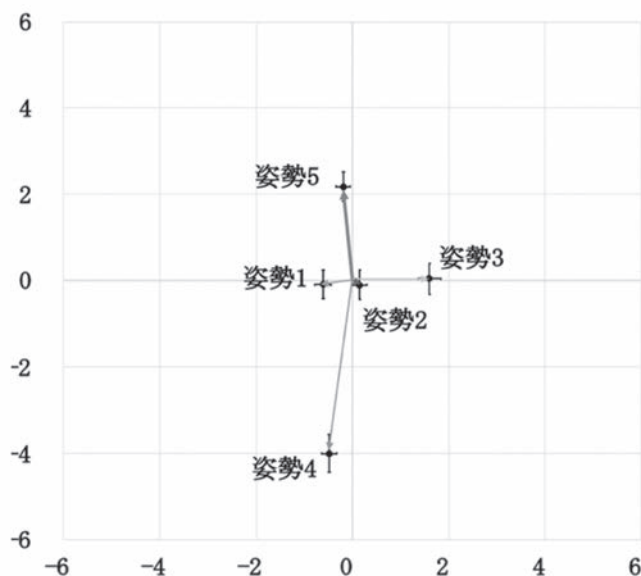


図2 5種類の姿勢の平均ズレ位置分布図  
(単位:mm)

## 凹線による方眼紙の升目において 升目の間隔と凹線と凸線との距離が触読性に及ぼす影響

### Influence of Grid Spacing and Distance between Concave and Convex Lines on Tactile Readability in Grids of Graph Paper with Concave Lines

南口 拓巳 (NANKO, Takumi) 指導：藤本 浩志

#### 1. 背景及び目的

特別支援教育（視覚障害）の現場では、点字教科書や点字図書が教材として広く用いられている。点字教科書や点字図書を校正する際には、原典である教科書や図書の内容に基づいて点訳が行われる。しかし、教科書や図書の中の図などを点図に置き換えて示す際には、凸状に盛り上がった凸点や凸線だけでは凸刺激が混在して触読し難いことから、点字教科書や点字図書の点図の中で方眼紙の升目には凹線が、湖などの空間領域には凹点パターンが多用されている。このような凹刺激は、凸点や凸線などの凸刺激の知覚を妨げずに学習上大切な情報を示すことができるため、非常に重要な刺激である。したがって、凸刺激よりも指先への刺激が少ない凹刺激を用いて適切に情報を提示するためには、その表示法に十分な配慮を行い、高い触読性を確保したものでなければならない。しかし、凹線や凹点パターンなどの凹刺激と触読性との関係を定量的に評価した学術的な知見は少ない。そのため、校正者が点図を校正する際には未だ経験則に依存する部分が大きく、校正の際に参考となる凹刺激の表示法に関する定量的な指針が求められている。

そこで本研究では、点字教科書や点字図書の点図の中でも方眼紙の升目を構成する凹線を想定し、凹線の凹点間隔、凹線間隔と触読性との関係を定量的に明らかにすることを目的とした。なお本実験は、早稲田大学の「人を対象とする研究に関する倫理委員会」の承認を得て行った。

#### 2. 方法

実験参加者として、若年晴眼者19名（平均年齢 $20.4 \pm 2.3$ 歳）を対象に、複数設けた凹線の中から凸線をなぞり、指定の凹線に指腹を置くことができるかを調べた。

提示刺激は、垂鉛版製版によって凸線を1本、凹線を7本並べたものを作成した。実験因子は、凹線と凸線との距離（2, 3, 4, 5, 6 [mm]）5条件を設定した。凹線と凸線の距離は点の中心間距離を表す。

実験参加者に対して縦方向に置かれた凹線と横方向に置かれた凹線に対し、触察方向として、左右方向と上下方向の2条件を用意した。全体の試行数は、提示刺激5条件に

つき、試行の終点（凹線1本目～6本目）6条件、触察方向2条件、凸線に対する凹線の向き2条件の計250であった。評価指標は正答率、回答時間、確信度を用いた。確信度は回答に対する確信度を5段階（1：確信なし～5：確信あり）で回答させた。

#### 3. 結果及び考察

正答率について、凹線と凸線との距離を要因とした分散分析を行った。そこで多重比較を行った結果、左右方向の触察において、2mmと4mm及び3mmと4mmの間に有意差がみられた。また、上下方向の触察において、2mmと4mm及び3mmと4mmの間に有意差がみられ、4mmと5mmの間にも有意差がみられた。図1に左右方向の触察における正答率の結果を示す。回答時間と確信度についても同様に、凹線間隔が狭いほど回答時間が長くなり、確信度が低くなる傾向がみられた。これは、凹線間隔が狭い条件では、触読する際に指が3本以上の凹線に触れるため、指先から得られる触覚情報が多く、本数を正確に数えることが困難であったためと考えられる。

#### 4. 結論

本研究では、点字教科書に用いられる凹点による方眼紙の升目を想定し、凹点間隔と凹線間隔が凹線を触読する際に及ぼす影響を定量的に評価することを目的とした。その結果、凹線間隔が狭いほど正答率が低く、回答時間が長くなり、確信度が低くなる傾向が見られた。

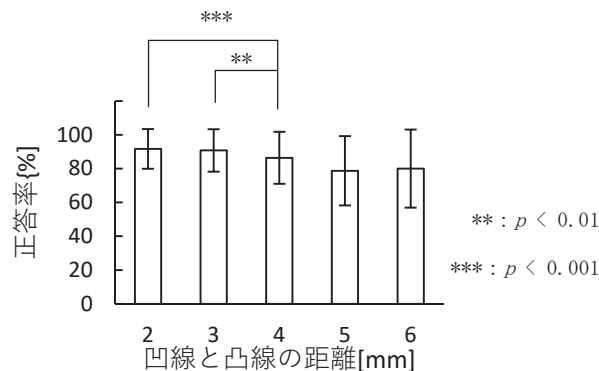


図1 左右方向の触察における正答率の結果

## マスク着用時の表情筋活動に関する研究

## A study on facial muscle activity while wearing a mask

藤田 夏菜子 (FUJITA, Kanako) 指導：宮崎 正己

## 1. 研究背景

2020年のコロナウイルス流行により、マスク着用が日常化している。厚生労働省の「新しい生活様式の実践例」の感染防止の3つの基本でも、マスク着用について言及がある [1]。しかし、マスクは表情の判断に大きく影響を与える口元を全て覆ってしまうため、相手に表情が伝わりづらくなる。すなわち、マスクによって顔が隠れることによって、他者の表情と一致した表情が非意図的に生じる表情同調が抑制され、表情筋活動に影響があるのではないかと予想した。本研究の目的は、マスク着用時とマスク未着用時の対話中の表情筋活動の違いと、伝達の質の違いを明らかにすることである。

## 2. 方法

本研究は、被験者同士で対話し、両者の表情筋電図を測定する実験1と、被験者と担当者で対話し、被験者のみ表情筋電図を測定する実験2の、2つの実験から構成されている。実験1では、19歳～24歳の14名を被験者とした。被験者は向かい合って椅子に座り、被験者同士で5分間の対話を行った。マスク着用同士・Aのみマスク着用・Bのみマスク着用・マスク未着用同士の4条件で対話を行った。実験2では、19歳～24歳の11名を被験者とした。被験者は担当者と向かい合って椅子に座り、担当者と5分間の対話を行った。マスク着用同士、マスク未着用同士の2条件で対話実験を行った。実験1, 2ともに、表情筋活動は左の大頬骨筋を、生体計測システムMPI60(BIOPAC Systems社)で計測した。また、主観評価として、各条件の対話終了毎にアンケートを実施した。

## 3. 結果

実験1では、一元配置分散分析とBonferroniの方法による多重比較を行った。その結果、表情筋活動に有意差はみられなかった。アンケートでは、相手のみマスク着用と比べてマスク未着用同士の方が、「聞き取りやすさ」では1%水準で、「対話のテンポ」「対話の温度感」「声の大きさ」では5%水準で有意に得点が高かった。

実験2ではウィルコクソンの符号付き順位検定を行った。その結果、マスク着用時と比べ、マスク未着用時の方が、大頬骨筋の筋活動が5%水準で有意に活発であった。アン

ケートでは、マスク着用時と比べてマスク未着用時の方が、「表情の読み取りやすさ」では1%水準で、「対話の楽しさ」「相手への信頼度」他8項目では5%水準で有意に得点が高かった。

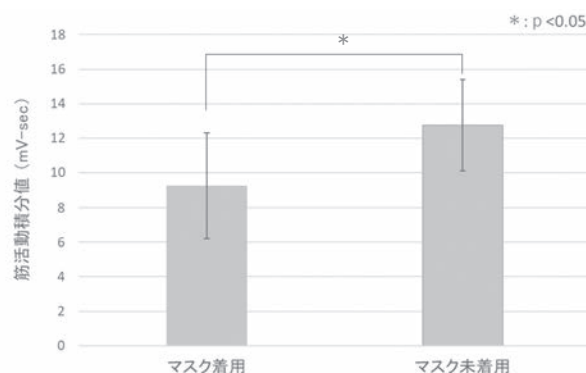


図1 実験2の5分間の大頬骨筋活動積分値

## 4. 考察

実験1において、表情筋活動に有意差がみられなかった要因として、自然な対話場面を想定した非構造化された実験という点で、統制が不十分であったことが考えられる。アンケート結果では、マスク未着用時の方が相手の声が聞き取りやすかったことから、対話のテンポが良く、対話の温度感が温かいと判断されたと考察する。また、実験2の大頬骨筋活動の結果は、表情同調によるものだと考察する。これは、アンケートの「表情の読み取りやすさ」について、マスク着用時と比べてマスク未着用時の方が、表情が読み取りやすかったという結果からも、表情同調が起こりやすくなり、表情筋が動いていたことが考えられる。また、表情同調による感情の追体験が起こったことで、「対話の楽しさ」や「相手への信頼度」などに影響を与えたと考えられる。マスク着用時の対話では表情が見えないことから、コミュニケーション時の齟齬が起こらないようにいつも以上に言葉に注意することが重要となるだろう。

## 参考文献

- [1] 厚生労働省, [https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431\\_newlifestyle.html](https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000121431_newlifestyle.html) (2021年12月24日閲覧)



# 機械学習によるソーシャルワーカーの非言語行動スキルの 対話ロボットへの実装

## Implementing Social Workers' Nonverbal Behavioral Skills into Interactive Robots Using Machine Learning

森木 海翔 (MORIKI, Kaito) 指導：菊池 英明

### 1. 背景と目的

現在日本は超高齢社会であり、高齢者の認知症有病率も年々上昇している。認知症予防の方法の一つに対話コミュニケーションがある。介護施設では介護士やソーシャルワーカーによる傾聴対話が行われているが、忙しさなどから十分な人員と時間の確保が難しい。そこで高齢者の話し相手になる対話ロボットの活躍が期待されている。

介護施設では介護士やソーシャルワーカーが高齢者とコミュニケーションを取る際に、身体動作を用いて大きく反応するような非言語行動スキルが用いられている。また、人間同士の対話であれば「大きく反応する」ことで、対話相手に「話を聞いてもらえるという実感を与える」ことも明らかになっている。しかし、対話ロボットの身体動作を用いて「大きく反応する」ことによって人間に与えられる効果は明らかになっていない。

そこで本研究では、対話ロボットが人間との対話中に、機械学習によって適切な場所で「大きく反応する」身体動作を行うことで「話を聞いてもらえるという実感を与える」ことができるか調査することを目的とする。

### 2. ソーシャルワーカーの非言語行動スキルの調査

ソーシャルワーカーが高齢者との対話で用いる具体的な「大きく反応する」という非言語行動スキルを、ソーシャルワーカーと認知症高齢者の対話映像を用いて明らかにした。その結果、ソーシャルワーカーが、高齢者の笑えるような話に対して上半身を大きく仰け反って笑う非言語行動スキルを用いていることがわかった。

### 3. WOZ法を用いた対話ロボットによる対話実験

対話映像から明らかになったソーシャルワーカーの上半身を大きく仰け反って笑う非言語行動スキルを、人間が笑う際の身体動作の特性を踏まえて対話ロボットに実装した。そして、非言語行動スキルが人間に対して有効に働くかを、ユーザから共有される笑える話にロボットが振る舞いを行うという場面にてWOZ法を使用した対話実験を行った。具体的には、非言語行動スキルの振る舞いの大きさと反応タイミングの効果について検証した。その結果、人間と対話ロボットの対話において、「大きく反応する」ことで相手

に「話を聞いてもらえるという実感を与える」度合いが増加するということが示唆された。また、被験者属性については、「ロボット否定的態度尺度」、「ロボット不安尺度」と「話を聞いてもらえるという実感」の度合いの評価値との相関分析から、ロボットへの否定的態度が高い人に対して、対話ロボットの大きな振る舞いがより有効であることが示唆された。さらに、ロボットの行動に対する不安感が強い人に対しても、対話ロボットの大きな振る舞いがより有効であることが示唆された。

### 4. 非言語行動スキルの機械学習による自動化及び雑談対話ロボットの構築

WOZ法を使用した実験で明らかになった人間に対して有効に働く身体動作について、名大会話コーパスを学習データとしてBERTのファインチューニングによるモデル構築を行い、人間との対話中に非言語行動スキルを使用できるように笑い判別モデルを構築した。また、NTTコミュニケーション科学基礎研究所のjapanese-dialog-transformersを用いて雑談対話が可能な応答生成モデルを使用し、自律対話を行うことができるロボットの構築を行った。

### 5. 自律対話を行う雑談対話ロボットによる対話実験

自律型対話ロボットを用いて、「大きな反応」を自動化した場合でも人間に対して有効に働くか実験を行った。その結果、対話中にロボットが人間の笑い話に対して自動で「大きく反応する」ことで、人間同士の対話同様に「話を聞いてもらえるという実感」の度合いが増加すること、及び「Agent's Attentiveness (あなたは、ロボットがあなたを意識し、気を配ってくれているように感じた)」の評価が高くなることが示唆された。被験者属性については、「ロボット不安尺度」と「Agent's Attentiveness」の相関分析を行った。その結果、「ロボットが会話中に的外れなことを話すのではないかと」といったロボットの会話能力に不安を抱えていたり、「ロボットに話しかけた時にどう答えたらよいのか」といったロボットとの対話への不安感が強い人に対して、笑い判別モデルを使用してロボットが対話中に「大きな反応」を行うことで「Agent's Attentiveness」の効果をもっと高めることが示唆された。