

自閉スペクトラム症の発話者コーパス作成と自動診断支援システム構築

崎下 雅仁¹ 小川 ちひろ¹ 土屋 賢治² 岩渕 俊樹²

岸本 泰士郎³ 狩野 芳伸¹

¹静岡大学 ²浜松医科大学 ³慶應義塾大学

{msakishita, cogawa}@kanolab.net, {tsuchiya, iwabuchi}@hama-med.ac.jp,
tkishimoto@keio.jp, kano@inf.shizuoka.ac.jp

1 はじめに

発達障害がある方の人口は近年増えつつあり、決して珍しくはない。2006年から2008年のアメリカ合衆国の調査では子供の6人に1人が発達障害を抱えていると明らかになった。これは全体の子供のうち15%に当たる(Boyle et al., 2011)。本稿では発達障害の一つである自閉スペクトラム症 (ASD)に着目する。Kanner (1943) がこの障害を autism と名付けた。ASD は対人的相互関係、対人的相互反応で用いられる非言語的コミュニケーション行動、および人間関係を発展・維持、および理解する能力などに障害があり、様々な状況における社会的コミュニケーションおよび対人的相互反応の持続的な障害によって特徴づけられる(American Psychiatric Association, 2013)。

ASD は診察者や診察時の環境の違い、患者の診察時の状態により、診断に揺れが出る可能性がある。このような問題を解決するために、重症度を定量化するための症状評価ツールが開発されているが、このツールを完全に自動化する研究は、我々が知る限り存在しない。本稿では発話特徴から ASD の重症度を自動推定する診断支援ツールを提案する。発話特徴について、ASD を持つ多くの方々は、発話のタイミング、文法、語彙、話す速度などの独特の特徴があり、知的障害および言語障害を伴う。

ASD がある方の発話に関する研究は ASD がない方の発話研究と比べて非常に少なく、ASD の日本語音声コーパスは公開されていない。過去の研究には言語特徴による ASD と定型発達を分類する研究がある (Rouhizadeh et al., 2014)。重症度を推定する研究では視線追跡 (Yaneva et al., 2018)、音声情報を使用するもの(Asgari et al., 2013)がある。

私たちは手作業による注釈付きの世界最大の日本語 ASD 発話者コーパスを作成した。本コーパスは、診断基準が臨床医によって ASD と診断された被験者と、インタビュアーの臨床心理士とのコミュニケ

ーションを記録した自閉症診断観察スケジュール (Autism Diagnostic Observation Schedule, ADOS) (Lord et al., 2000) 評価に基づいている。ADOS は、ASD 重症度得点を割り当てることができる基準の1つである。ADOS コーパスを使用して、自動重症度評価ツールを実装し、どの特徴量が重症度特定に有効かを分析した。

私たちの研究は、言語聴覚士による言語リハビリテーション、ASD の可能性があるのかチェックをするスクリーニングを自動でサポートするための基礎となる可能性があることを示唆している (Sakishita et al., 2019)。

2 ADOS および ADOS コーパス

2.1. Autism Diagnostic Observation Schedule (ADOS)

ADOS は ASD 診断の基準の1つであり、ASD 疑いのある被験者について、コミュニケーション、相互対人関係、遊び/想像力、および限定的/反復行動を評価することができる、標準化された半構造化評価である (Lord et al., 2000)。ADOS は研究目的を可能にする特別な ADOS ライセンスを有する臨床心理士によってのみ実施することができる。ADOS には4つのモジュールがあり、発達レベルと対象年齢、会話の流暢さに応じてモジュールを選択できる。1 から4までのモジュールの中で、第4モジュールは最も発達レベルが高く、会話が流暢な成人を対象としている。ADOS を行う臨床心理士は、ADOS のプロトコルに従って各タスクを実行し、観察された行動を評価し、ADOS アルゴリズム (Lord et al., 2000) に従って被験者を評価する。ADOS 得点と ASD 重症度の間には一定の相関関係があることがわかっている (Gotham et al., 2009)。一回の施行の所要時間は約40~60分である。

表 1. 本稿で利用した ADOS 得点

| カテゴリ | 得点名 | 説明 | 範囲 |
|---------|------------------------------|---|------|
| 意思伝達 | <i>STER</i> | 単語や短文の常同的な使用 | 0-2 |
| | <i>CONV</i> | 会話 | 0-2 |
| | <i>DGES</i> | 叙述的、慣習的、道具的、情報提供的な身振り | 0-2 |
| | <i>EGES</i> | 協調的、あるいは感情的な身振り | 0-2 |
| | <i>Communication (total)</i> | 「意思伝達」の小計 | 0-8 |
| 相互的対人関係 | <i>EYE</i> | 普通と異なるアイコンタクト | 0,2 |
| | <i>EXPO</i> | 他者に向けられた顔の表情 | 0-2 |
| | <i>EMO</i> | 他者の感情についての共感/コメント | 0-2 |
| | <i>RESP</i> | 責任 | 0-2 |
| | <i>QSOV</i> | 対人的働きかけの質 | 0-2 |
| | <i>QSR</i> | 相互的反応の質 | 0-2 |
| | <i>ARSC</i> | 相互的な対人的意思伝達の量 | 0-2 |
| | <i>Interaction (total)</i> | 「相互的対人関係」の小計 | 0-14 |
| | <i>Total</i> | <i>Communication (total)</i> + <i>Interaction (total)</i> | 0-22 |

2.2. ADOS コーパス

我々の ADOS コーパスは、浜松医科大学で収録された ADOS の映像記録を元に作成された。我々は発話に焦点を当てているので、映像内の音声データのみを使用している。同様の理由で、言語解析ができるよう、複雑な会話を含むことができる第 4 モジュールを使用する。表 1 は、第 4 のモジュールの ADOS 得点の種類をまとめたものである。3 つの合計得点 (*Communication (total)*、*Interaction (total)*、*Total*) のいずれかが閾値を超えると、被験者は ASD として診断される。得点が高いほど、ASD の重症度は高くなる。同じ臨床心理士がすべての ADOS を記録し、実施した。この臨床心理士は研究目的で ADOS を使用できる公式の臨床心理士として登録されている。

我々のコーパス中の被験者は全て、精神科医によって ASD と診断されているが、精神科医は他の多くの基準を使用しているため、被験者が必ずしも ADOS 閾値を上回るとは限らない。

我々は、ADOS の録音音声を手書き起こした。千葉大学 3 人会話コーパス¹のアノテーション基準に基づいて、アノテーションを追加した。17 種類のアノテーション (休止、引き伸ばし、詰まり、語の中断、上昇調、フィラー、応答形感動詞、言いさし_意図された語が同定可能、言いさし_意図された語が不明、いい誤り、漢字表記できなくなった文字、仮名、歌、聞き取れない言語音、笑い、呼吸、読み仮名) を使用した。

ADOS コーパスには「パズルのおもちゃ組立て」部分、「テキストの無い絵本のストーリーの説明」部分、「リポートエリアの絵の叙述」部分を使用した。これらの部分には、インタビュアーの臨床心理士と被験者との言語コミュニケーションが含まれている。合計被験者数は 32 人 (男性 25 人、女性 7 人) で合計時間は 560 分である。被験者の年齢幅は 17 ~ 55 歳である。アノテーターの数は 2 人である。各書き起こしとアノテーションは一人のアノテーターによって行われ、すべてのファイルは間違いを訂正するため二重チェックした。

3 ADOS 得点予測

前節で紹介した ADOS コーパスを使用して、すべての被験者の ADOS 得点を予測する。表 2 は、訓練に使用した特徴量を示している。特徴量を被験者のプロフィール特徴量、言語特徴量、そして非言語特徴量に分類した。言語および非言語特徴量には、合計数および平均数 (録音時間の長さで除算) が含まれる。言語特徴量は、形態素解析器 JUMAN (Kurohashi et al., 2009) を用いて出現回数および語彙数として抽出した。その他の特徴量はアノテーションから抽出した。予測結果は 5 分割交差検証によって得た。さらに、ADOS 得点の予測に役立つと見なされる特徴量を、感度分析 (Tanabe et al., 2013) を用いて選択した。

¹ <http://research.nii.ac.jp/src/Chiba3Party.html>

表 2. 訓練に使用した特徴量。言語、非言語特徴量は1秒あたりの数もこれに加わる。

| カテゴリ | 特徴量 | |
|--------|--------|---------|
| プロフィール | 性別 | 年齢 |
| | 形態素 | 語彙 |
| | 内容語 | 内容語語彙 |
| | 名詞 | 名詞語彙 |
| | 形容詞 | 形容詞語彙 |
| | 動詞 | 動詞語彙 |
| | 副詞 | 副詞語彙 |
| | 助詞 | 助詞語彙 |
| | 助詞 “が” | 助詞 “に” |
| | 助詞 “は” | 助詞 “を” |
| | 助詞 “も” | 助詞 “で” |
| | 助詞 “と” | 6文字以上単語 |
| | 指示語 | 指示語語彙 |
| | 接続詞 | 接続詞語彙 |
| | 否定語 | 疑問 |
| | 感動詞 | |
| 非言語 | 合計時間 | 発話時間割合 |
| | 反応時間 | フィラー |
| | 笑い | 言いさし |
| | 言い間違い | 不明瞭語 |
| | | |

3.1. 予測結果と議論

臨床心理士が ADOS の資格を得るためには、正解と完全一致する得点を割り当てるのが求められる。ADOS 得点は整数であるが、少数値を出力するシステムに当てはめると、合計得点項目以外の得点の誤差が 0.5 より小さく (*EYE* の場合は 1.0) すれば完全に正答し ADOS 資格の要求を満たしうる。この誤差を許容誤差と呼ぶことにする。例えば、システムが *STER* に対して正解得点が 0 であるとき、0.2 を出力したと仮定する。*STER* は整数値 0、1 または 2 のいずれかでなければならない。システム出力値に最も近い整数値が正解得点の場合、出力値は正しいと見なすことができる。この例では、システムの出力は 0.2 で、最も近い整数値は 0 である。0 が正解得点であるため、この出力は正しいと見なすことができる。合計点得点項目についての許容誤差も同様に、小計項目の許容誤差の合計とする。*Communication (total)*、*Interaction (total)*、*Total* の許容誤差はそれぞれ 2.0、4.0、および 6.0 となる。以下この評価基準を採用する。

表 3 に予測結果を示す。結果許容誤差が 0.5 の RMSE (平均二乗誤差の平方根) は、0.66 (*STER*)、0.74 (*CONV*)、0.48 (*DGES*)、0.87 (*EGES*)、0.62 (*EXPO*)、0.62 (*EMO*)、0.83 (*RESP*)、0.53 (*QSOV*)、0.41 (*QSR*) および 0.81 (*ARSC*) であった。許容誤差が 1.0 の *EYE* の RMSE は 0.68 であった。これらの得点項目の RMSE

表 3. ADOS 得点の SVR による予測結果

| 得点名 | RMSE |
|------------------------------|------|
| <i>STER</i> | 0.66 |
| <i>CONV</i> | 0.74 |
| <i>DGES</i> | 0.48 |
| <i>EGES</i> | 0.87 |
| <i>Communication (total)</i> | 1.56 |
| <i>EYE</i> | 0.68 |
| <i>EXPO</i> | 0.62 |
| <i>EMO</i> | 0.62 |
| <i>RESP</i> | 0.83 |
| <i>QSOV</i> | 0.53 |
| <i>QSR</i> | 0.41 |
| <i>ARSC</i> | 0.81 |
| <i>Interaction (total)</i> | 2.40 |
| <i>Total</i> | 3.53 |

はすべて許容誤差を下回っているかそれに近い値となった。合計得点項目について RMSE は 1.56 (*Communication (total)*)、2.40 (*Interaction (total)*) および 3.53 (*Total*) であった。合計得点項目についてシステムの誤差はすべて許容誤差を下回っている。

臨床心理士は、ADOS テストの全体を通して ADOS 得点を割り当てる必要がある。しかし我々の結果は、ADOS テストの一部のみであっても得点を予測できることを示唆している。

3.2. ADOS 得点と特徴量の相関係数

重症度と関連が深い特徴量を調べるために、ADOS 得点と特徴量との相関係数を計算した。表 4 は、上位 2 つの正および負の相関係数が示されている。最初に注目する特徴量は *指示語語彙* である。表 4 中の ADOS 得点 4 項目の負の相関に *指示語語彙* または *指示語語彙/s* が見られる。被験者が自身の状況や場所を正確に捉え、伝えたい事物を臨床心理士に指示語を適切に用いて伝えることができるが、頻繁に使用しない場合、重大度は低くなることを意味している。

次に注目した特徴量は笑いである。表 4 中の 5 つの ADOS 得点項目の負の相関に笑いまたは *笑い/s* が見られる。笑いの数が多いほど、ADOS 得点は低くなる傾向があることがわかった。ADOS は本来笑いを測定することは想定されていないため、これは新しい発見である。

次に注目した特徴量は *不明瞭語* である。表 4 の 9 つの ADOS 得点項目の正の相関に、*不明瞭語* または *不明瞭語/s* が見られる。これは、聞き手が認識できない語が多ければ多いほど重症度が高いことを意味する。

表 4. 全ての特徴量と ADOS 得点との相関係数。“/s”は録音時間で除算した数を表す。

| STER | | CONV | | DGES | | EGES | |
|-----------------------|--------|----------|--------|--------|--------|----------|--------|
| 笑い | -0.360 | 助詞語彙 | -0.642 | 指示語語彙 | -0.661 | 年齢 | -0.391 |
| 笑い/s | -0.358 | 語彙 | -0.632 | 合計時間 | -0.598 | 指示語語彙 | -0.339 |
| 言いさし | 0.416 | 感動詞/s | 0.434 | 名詞語彙/s | 0.493 | 動詞語彙/s | 0.19 |
| 不明瞭語 | 0.422 | 不明瞭語/s | 0.44 | 語彙/s | 0.493 | 不明瞭語/s | 0.208 |
| Communication (total) | | EYE | | EXPO | | EMO | |
| 指示語語彙 | -0.614 | 笑い/s | -0.348 | 動詞 | -0.426 | 助詞 “が”/s | -0.478 |
| 助詞語彙 | -0.609 | 性別 | -0.327 | 合計時間 | -0.425 | 助詞 “を”/s | -0.361 |
| 発話時間割合/s | 0.444 | 接続詞/s | 0.353 | 助詞語彙/s | 0.419 | 指示語/s | 0.339 |
| 不明瞭語/s | 0.572 | 不明瞭語/s | 0.418 | 不明瞭語/s | 0.517 | 感動詞/s | 0.384 |
| RESP | | QSOV | | QSR | | ARSC | |
| 性別 | -0.41 | 疑問 | -0.457 | 笑い | -0.653 | 反応時間/s | -0.608 |
| 年齢 | -0.385 | 助詞 “で”/s | -0.389 | 笑い/s | -0.574 | 語彙 | -0.574 |
| 名詞語彙/s | 0.38 | 反応時間 | 0.338 | 反応時間 | 0.328 | 感動詞/s | 0.392 |
| 言い間違い/s | 0.384 | 助詞 “は”/s | 0.342 | 不明瞭語/s | 0.405 | 助詞語彙/s | 0.518 |
| Interaction (total) | | Total | | | | | |
| 笑い | -0.508 | 指示語語彙 | -0.586 | | | | |
| 笑い/s | -0.506 | 笑い | -0.553 | | | | |
| 助詞語彙/s | 0.417 | 助詞語彙/s | 0.404 | | | | |
| 不明瞭語/s | 0.538 | 不明瞭語/s | 0.591 | | | | |

4. 終わりに

ASD の診断を下すための ADOS 音声記録を用いて、アノテーション付き ASD 発話者コーパスを構築した。我々の ADOS コーパスは、ASD の日本語話者のコーパスとして世界最大である。このコーパスを用いて機械学習手法により ADOS 得点 (ASD の重症度) を予測した。予測結果は ADOS の専門家とほぼ同程度の誤差水準を達成した。ADOS 得点と特徴との間の相関係数は、指示語語彙、笑い、および不明瞭語特徴が ADOS 評価に影響を与えることを示した。今後の課題として、実際の臨床現場で役立つよう、子供の発話分析や音声認識のために、ADOS コーパスのサイズを増やすことを計画している。

謝辞

本研究は JST AIP-PRISM 助成金番号 JPMJCR18Z7、AIP チャレンジ、JST CREST、科学研究費補助金の援助を受けて行なわれました。被験者の方々、そして ADOS を実施していただいた臨床心理士の松本かおり先生に心からお礼を申し上げます。

◇ 参考文献 ◇

[1] American Psychiatric Association. 2013. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). American Psychiatric Publishing.

[2] Asgari, M., Bayestehtashk, A. and Shafran, I. (2013). Robust and accurate features for detecting and diagnosing autism spectrum disorders. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH (pp. 191–194).

[3] Boyle, C. a, Boulet, S., Schieve, L. a, Cohen, R. a, Blumberg, S. J., Yeargin-Allsopp, M., Visser, S. and

Kogan, M. D. (2011). Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics*, 127(6), 1034–1042. doi:10.1542/peds.2010-2989

[4] Gotham, K., Pickles, A. and Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(5), 693–705. doi:10.1007/s10803-008-0674-3

[5] Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*. doi:10.1105/tpc.11.5.949

[6] Kurohashi, S. and Kawahara, D. (2009). Japanese Morphological Analysis System JUMAN 6.0 Users Manual. Retrieved from

[7] Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, and Rutter, M. (2000). Autism Diagnostic Observation Schedule (ADOS). *Journal of Autism and Developmental Disorders*. doi:10.1007/BF02211841.

[8] Rouhizadeh, M., Prud'hommeaux, E., van Santen, J. and Sproat, R. (2014). Detecting linguistic idiosyncratic interests in autism using distributional semantic models. *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 46–50. Retrieved from <http://www.aclweb.org/anthology/W/W14/W14-3206>

[9] Sakishita, M., Ogawa, C., Tsuchiya, K. J., Iwabuchi, T., Kishimoto, T., Kano, Y. (2019). Autism Spectrum Disorder's Severity Prediction Model Using Utterance Features for Automatic Diagnosis Support. *Proceedings of AAAI-19 Joint Workshop on Health Intelligence, W3PHIAI'19*. (accepted)

[10] Tanabe, K., Kurita, T., Nishida, K., Lučić, B., Amić, D. and Suzuki, T. (2013). Improvement of carcinogenicity prediction performances based on sensitivity analysis in variable selection of SVM models. *SAR and QSAR in Environmental Research*, 24(7), 565–580. doi:10.1080/1062936X.2012.762425

[11] Yaneva, V., Ha, L. A., Eraslan, S., Yesilada, Y. and Mitkov, R. (2018). Detecting Autism Based on Eye-Tracking Data from Web Searching Tasks. *Proceedings of the 18th Web for All Conference on The Internet of Accessible Things, W4A 2018*, (di).