

## 自閉症スペクトラム児と保護者間のインタラクション分析

## Interaction Analysis between Children with Autism Spectrum and Their Parents

田中 宏季  
Hiroki Tanakaサクリアニ サクティ  
Sakriani Saktiグラム ニュービッグ  
Graham Neubig戸田 智基  
Tomoki Toda中村 哲  
Satoshi Nakamura

奈良先端科学技術大学院大学 情報科学研究科

Graduate School of Information Science, Nara Institute of Science and Technology

Autism spectrum disorders are developmental disorder characterised as social and communication, and it affects verbal and non-verbal expressions. Previous works report differentiation of children with and without autism spectrum disorders. In this paper, we analysed several linguistic and prosodic cues, and pauses before new turns in child-parent interaction of narrative. We found that our method is useful to classify autism and typical development, and there is a small data differences in relationship of pauses and linguistic cues. This result is able to apply automatic dialogue aid system for social skills training.

## 1. はじめに

自閉症スペクトラム障害 (ASD) とは先天的な脳機能の発達障害であり, 1943 年にレオカナー, 1944 年にハンスアスペルガーによって, ほぼ同時期に初めて報告された [Kanner 43, Asperger 44]. 現在アメリカ精神医学会の定める ASD の特性として, 1) 定性的な社会性の障害とコミュニケーションの障害, 2) こだわりと想像力の障害, の2つをあげている [APA 13]. 特に1) 社会性とコミュニケーションの障害に関しては, ASD の中心となる障害であると言われている. 社会性とコミュニケーションの重要な概念として共感が挙げられる. 共感とは, 認知と表出の2つの要素に分解することができる. 認知とは心の理論課題あるいは感情認識能力などであり, 表出とは認知をした後, それに対して適切な応答を行う能力である [Davis 83]. これら2つの能力によって, 人間の基本的な社会的コミュニケーションが成立する. 共感能力の度合いを定量的に測定する事により, ASD の特性理解, スクリーニングへの応用, また治療法・教育法の方針策定が可能になると考えられる [Heeman 10].

ASD 者と定型発達 (TD) 者の違いを明らかにする研究は幾つか行われており, 共感においては, 発話応答時間とフィラーの使用に関するもの [Heeman 10], 韻律に関するもの [McCann 03], 声質に関するもの [Asgari 13], 予期しない語彙の使用に関するもの [Rouhizadeh 13] などがあげられる. しかしながら, これらの情報について包括的に分析し, スクリーニングへの応用に向けたシステム開発についてはこれまで言及されていない. 本研究ではこれらの違いを包括的に分析し, ASD 児のスクリーニングシステム開発に応用する.

本研究は児童の9歳から13歳までを対象として行う. 使用するデータに関しては, 児童において物語発話 (ナラティブ) を使用する有効性が報告されているため [Davis 04], 本研究では児童における自らの印象に残った体験の発話を題材に, 物語発話における保護者とのインタラクション中の特性分析, および米国のデータとの差について検討を行う.

## 2. データ説明

本研究では子どもと保護者間のインタラクションデータを使用する. データは, 粘土を使用した自由なごっこ遊び, ゲーム (ジェンガ), 物語発話, 自然対話の4つのセッションで区切られており, それぞれ10分の計40分である. インタラクション中は, ビンマイクとビデオカメラにより, 子どもと母親の音声と動画をそれぞれ記録している. 本研究では, パイロット実験として ASD 児4名 (男児:3名, 女児:1名) と TD 児2名 (男児:1名, 女児:1名) の物語発話のデータを使用する. 知能指数 (IQ) は全員70のカットオフ値以上である. 被験者の情報を表1に示す. 物語発話は, 「これまで印象に残っている体験」について, 子どもが, 母親がそれぞれ5分主導で話し, 聞き手がそれに対して質問するという内容となっている. 5分時間が経過すると, データ収録者によりターン交代の合図がなされ, 主導者と聞き手の役割が交代する.

表1: 被験者の情報.

Subject	S1	S2	S3	S4	S5	S6
Age	10	10	10	13	10	12
Outcome	ASD	ASD	ASD	ASD	TD	TD

その内, 子どもが主導者, 母親が聞き手である5分間のデータを使用した. 子どもと母親の発話は USC Rachel コーパス [Mower 11] の定義に従い, 書き起こしがなされた. ここでは, 1秒以上発話間のポーズがあった際に1発話として区切る. ASD と TD での発話を116ずつグループ内でランダムに抽出し使用した.

## 3. インタラクション分析

本節は, データから ASD 児と TD 児のインタラクション中における違いを明らかにすることを目的としている. 最後は特徴量を抽出の上, ASD 児と TD 児の自動識別を行う.

## 3.1 発話応答時間

Heeman らにより, 自然対話において ASD 児では TD 児に比べ質問への応答が遅れるという結果が得られている

連絡先: 田中 宏季, 奈良先端科学技術大学院大学, 奈良県生駒市高山町 8916-5, hiroki-tan@is.naist.jp

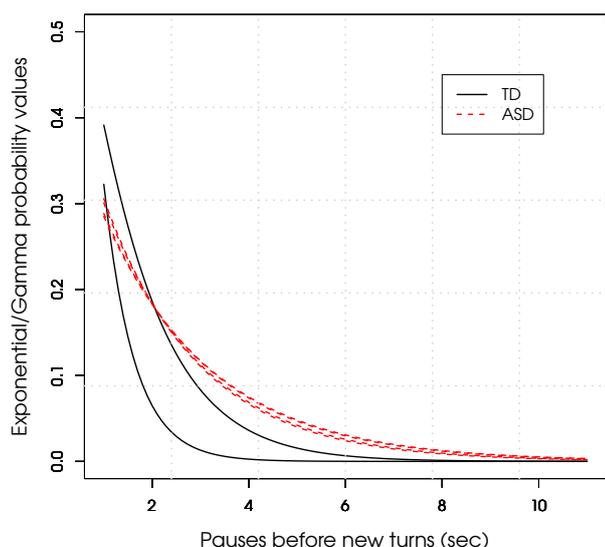


図 1: 発話応答時間に対するガンマ/指数分布によるフィッティング。ASD 児と TD 児に対し最尤推定法 (MLE) でパラメータを計算。

[Heeman 10]. 本研究では物語発話でも同様の傾向があるかを調査した。発話応答時間 (Pauses before new turns) 値を、保護者が発話をした後、児童が応答するまでの時間により算出した。ここで発話のオーバーラップは除外している。

我々は発話応答時間に対して、ガンマ分布あるいは指数分布についての当てはまりの良さを検定を行った (ここで指数分布はガンマ分布で形状母数を 1 とした特殊なケースである)。図 1 はガンマ分布あるいは指数分布を発話応答時間にフィッティングした分布のグラフである。ガンマ分布、指数分布の内、コルモゴロフ-スミルノフ検定によりより良くフィットしている分布を採用した。結果として TD 児の 1 名のみガンマ分布によりよく当てはまり、残りの 5 名は指数分布によりよく当てはまった。ここでは、どの被験者においても  $p > 0.6$  となり、十分に当てはまっていると解釈することが出来る。図 1 の結果として ASD 児では TD 児と比較して発話応答時間が遅れる傾向があることが確認された。我々は指数分布の期待値を特徴量として算出した。

また我々は、発話応答時間が直前の保護者の質問の種類と関連があるかも調査した。開いた質問 (open-question) と閉じた質問 (closed-question)、および質問以外 (non-question) の 3 種類をラベル付けし、それぞれに対する発話応答時間の平均値を算出した。結果として、ASD 児の方が TD 児に比べ、質問の種類に関係なく発話応答時間が遅れることが確認された。また質問の種類による違いでは、開いた質問と閉じた質問では差がなく、質問と質問以外で差が見られることを確認した。

### 3.2 Words per minute

ASD 児と TD 児でそれぞれ Words per minute (WPM) を算出した。約 5 分間のデータを使用した為、5 回分の WPM の平均値を特徴として算出した。ここでの単語とは Mecab<sup>\*1</sup>を

\*1 <https://code.google.com/p/mecab/>

表 2: 質問の種類と発話応答時間の関係性。平均値と標準偏差を示している。

Question type	TD	ASD
Closed-question	0.47 (0.46)	1.61 (1.87)
Open-question	0.43 (0.34)	1.76 (1.51)
Non-question	0.95 (1.18)	2.60 (3.64)

使用した際の、1 形態素と定義している。結果を表 3 に示す。結果として、ASD 児と TD 児での有意差は見られなかったが、個人差が見られる事がわかった。これらは外交的・内向的ななどの性格尺度とも関連していると考察される [Mairesse 07]。

表 3: WPM の平均値。

Subj.	Averaged WPM
S1	18.25
S2	86.75
S3	23.75
S4	115.5
S5	99.25
S6	103.5

### 3.3 予期しない語彙

予期しない語彙の使用について、TF-IDF と対数 odds の 2 指標を用いて評価を行った。TF-IDF とは tf (単語の出現頻度) と idf (逆文書頻度) の 2 つの指標にもとづいて計算される。我々は、以下の式により ASD 児と TD 児の TF-IDF 値を計算した。ここで、児童の物語発話番号を  $i$ 、物語発話中の単語番号を  $j$  としている。 $c_{ij}$  は、物語発話  $i$  中の単語  $j$  の出現回数である。 $f_j$  は単語  $j$  を含んでいる物語発話の数であり、 $D$  は全体の物語発話の個数である。

$$tf - idf_{ij} = (1 + \log c_{ij}) \log \frac{D}{f_j}$$

対数 odds ratio は TF-IDF とは別の尺度であり、情報検索あるいは抽出タスクなどで使用される。これは物語発話  $i$  中に出現する特定の語  $j$  のオッズを表している。特定の単語が特定の物語発話中に表れる確率を  $p_1$  とし、その単語がその他の物語発話中に表れる確率を  $p_2$  とした際、対数 odds ratio は以下の式となる。

$$\text{odds ratio} = \frac{\text{odds}(p_1)}{\text{odds}(p_2)} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$$

TF-IDF と対数 log odds ratio は特定の単語  $j$  が物語発話  $i$  に特質するものかどうかを表しており、これにより予期しない語彙の使用を推測することが出来る。我々は、これらの 2 つの尺度を算出し、さらに既存のコーパスデータ量が少ないため、外部資源である日本語 Wordnet<sup>\*2</sup>によるフィルタ処理を行い、Wordnet に存在する単語については値を算出しない手法をとった [Rouhizadeh 13]。これらの特徴量についての結果を表 4 に示す。これより今回使用した物語発話では、先行研究 [Rouhizadeh 13] のような ASD 児と TD 児での差が生じ

\*2 <http://www.omomimi.com/wjpn/>

る結果を得られていないことがわかる。これは本研究では物語発話の繰り返し（リテリング）など制約のあるタスクを使用していないため、使用する語彙に個人差が生じたためだと考えられる。

表 4: TF-IDF と対数 odds ratio 値。W.N. は Wordnet の略。

Subj.	TF-IDF+W.N.	Log-odds+W.N.	Sum (TF-IDF, Log-odds)+W.N.
S1	0.50	1.01	1.52
S2	0.58	0.49	1.08
S3	0.66	1.23	1.89
S4	0.66	0.31	0.96
S5	0.74	0.49	1.23
S6	0.62	0.44	1.06

### 3.4 言語・韻律・声質特徴量

先行研究 [Tanaka 14] において、言語と韻律特徴量が ASD 児と TD 児を識別するのに有効であることが報告されているため、本研究でも物語発話全体に対して同様の特徴量を抽出した。本研究では先行研究により有意差が確認された 7 種類の特徴量（6 文字以上の単語の頻度、社交語の割合、感情語の割合、認知語の割合、同意の割合、フィラーの割合、基本周波数の標準偏差）に加え、声質特徴量として ASD 児と TD 児で有意差のあった h1a3 [Hansen 95] も算出した。1 発話毎にそれぞれの特徴量を抽出した上、全体の平均値を算出し特徴量とした。

### 3.5 ASD 児と TD 児の識別

少ないサンプル数 (N=6) での識別であるため、K=1 において K 近傍法を使用し ASD 児と TD 児の物語発話の識別を行った。特徴量としては本節で述べたもの全て（発話応答時間（指数分布の期待値）、WPM, TF-IDF, 対数 odds ratio, 6 文字以上の単語の頻度、社交語、感情語、認知語の割合、同意、フィラーの割合、基本周波数の標準偏差、h1a3）を使用し、leave-one-speaker-out 交差検定を行った。結果として全話者を正しいクラスに識別できた。よってこれらの特徴量は ASD 児と TD 児を識別するのに有効であることが確認された。

## 4. データ比較

本節では ASD 児の発話応答時間に対して、コーパスの差について検討する。

表 5: 使用した言語特徴量。

カテゴリ	説明
一般記述	1 発話中の単語数 (WPS) 6 文字以上の単語、笑いの頻度
文構造	代名詞、接続詞、否定 数量詞、数値の割合
心理プロセス語	社交語、感情語、認知語、知覚語 生理語、関係語の割合
個人的関心	仕事、達成、レジャー、家庭の割合

比較用データとして、USC Rachel コーパス [Mower 11] (米国の ASD 児童 9 名を対象) を使用する。USC Rachel コーパスを使用し、発話応答時間と、ASD 児と保護者の言語情報との関連性を調べている研究が存在する [Theodora 13]。本研究

では、上記の研究と同じ枠組みでの比較調査を行った。なお、USC Rachel コーパスも、児童-保護者インタラクションのセッションを含んでおり、書き起こしの定義は本研究と同様である。3.1 節の発話応答時間に対して 70%クォンタイル値を閾値として Short pause (S) と Long pause (L) の 2 種類のグループに分類した。これらの関係と児童の当該発話、保護者の直前の発話それぞれに対して関連性を分析した。ここで抽出した言語特徴量を表 5 にまとめる。言語特徴量は先行研究 [Tanaka 14] に従い、Linguistic Inquiry and Word Count (LIWC) 辞書を用いて抽出した。

表 6, 7 に、USC Rachel コーパスと日本語コーパスにおける、ブートストラップ法でリサンプリングした言語特徴量の平均値が有意に大きい発話応答時間を示している。2 つを比較すると、ほぼ一致した傾向が見られることを確認できるが、WPS に対しては異なる結果を示していることも確認できる。

表 6: USC Rachel コーパスにおける ASD 児と保護者の発話と、発話応答時間の関係性。

Subj.	Child				Parent		
	WPS	conj.	affect	nonflu.	adverb	cogn.	percept.
S1	L*	L*	S*	-	L*	L*	L*
S2	L*	L*	S†	L*	L*	L*	L*
S3	L*	L†	-	S†	L*	L*	L*
S4	-	-	-	L*	L*	L*	L*
S5	L†	-	-	-	L*	L*	L*
S6	L*	-	S*	-	L*	L*	-
S7	L†	-	S†	-	L†	-	-
S8	L*	-	-	-	L*	L*	L*
S9	-	-	-	S†	L*	L*	L*

表 7: 日本語コーパスにおける ASD 児と保護者の発話と、発話応答時間の関係性。

Subj.	Child				Parent		
	WPS	conj.	affect	nonflu.	adverb	cogn.	percept.
S1	S*	-	-	-	S*	L*	-
S2	S†	-	S*	-	L*	L*	L*
S3	S†	-	-	-	L*	L*	L*
S4	S*	-	-	-	-	-	-

日本語コーパスにおいて、WPS が Short pause で多くなる傾向に対し、調査を行った。逆となる Long pause の発話に対して、1 語のみで応答する傾向が高くなることが分析により確認された為、その 1 語について LIWC を用いた。図 2 に結果を示しており、これより、例えば S1 では長く発話応答時間がかかる際、同意語 1 語を使用する傾向があることがわかった。

## 5. まとめ

本研究は、ASD 児と保護者間でのインタラクションにおける特性に着目している。物語発話のデータを使用し、ASD 児では TD 児と比べ保護者の発話に対する発話応答時間が遅れる傾向があることを示した。またこれらは質問の種類に関わらず ASD 児共通の特性だと確認された。さらにデータの比較についても言及しており、日本語コーパスでは、発話応答時間が長くなる発話の特徴として、1 語あるいは少ない単語数である割合が高くなることもわかった。これには、共感の内、表出面だけでなく、保護者の言ったことを認知するという過程も含まれているため、今後はこれら表出と認知の両側面から検討をし

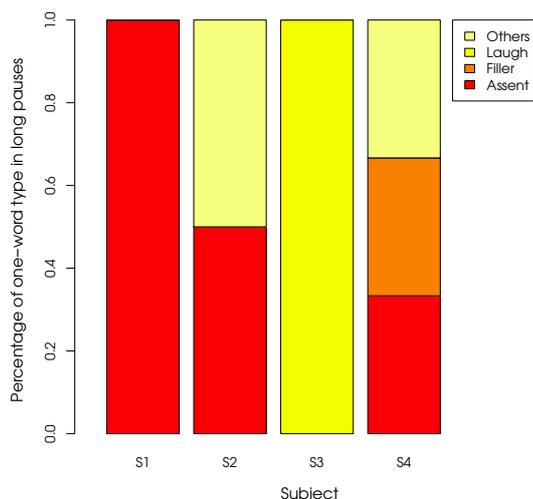


図 2: Long pause における 1 語発話の言語特徴量.

ていく必要がある。さらに、本研究の枠組みを音声認識を含め自動化し、対話システムにより ASD を自動識別するモデルを開発し、スクリーニングシステムにつなげていくことが必要である。

## 6. 謝辞

本研究に参加してくれた子どもと保護者に感謝を申し上げます。奈良教育大学の岩坂英巳先生には貴重なご意見とご協力をいただき、ここに感謝を申し上げます。本研究は Signal Analysis and Interpretation Laboratory (SAIL), USC のインターンシップで行われた結果をまとめたものである。本研究の一部は JSPS 科研費 24240032 の助成を受け実施したものである。

## 参考文献

- [Kanner 43] Kanner L. Autistic disturbances of affective contact. *Nervous Child* 2, 217-250 (1943).
- [Asperger 44] Asperger, H. Die „Autistischen Psychopathen“ im Kindesalter. *European Archives of Psychiatry and Clinical Neuroscience* 117, 76-136 (1944).
- [APA 13] American Psychiatric Association. *The diagnostic and statistical manual of mental disorders*, 5. Washington, D.C. (2013).
- [Davis 83] Davis M. H., Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology* 44, 113 (1983).
- [Heeman 10] Heeman, P. A., Lunsford, R., Selfridge, E., Black, L. & Van Santen, J. Autism and interactional aspects of dialogue. *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 249-252 (2010).
- [McCann 03] McCann, J. & Peppe, S. Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*, 325-350 (2003).
- [Asgari 13] Asgari, M., Bayestehtashk, A. & Shafran, I. Robust and Accurate Features for Detecting and Diagnosing Autism Spectrum Disorders, 191-194 (2013).
- [Rouhizadeh 13] Rouhizadeh, M., Prud'hommeaux, E., Roark, B. & van Santen, J. Distributional semantic models for the evaluation of disordered language. *Proceedings of NAACL-HLT*, 709-714 (2013).
- [Davis 04] Davis, M., Dautenhahn, K., Nehaniv, C. L. & Powell, S. D. Towards an Interactive System Facilitating Therapeutic Narrative Elicitation in Autism, *Third International Conference on Narrative and Interactive Learning Environments (NILE 2004)*, Edinburgh, Scotland. (2004).
- [Mower 11] E. Mower, M. Black, E. Flores, M. Williams, & S. Narayanan, Rachel: Design of an emotionally targeted interactive agent for children with autism, *ICME, Barcelona, Spain*, 1-6 (2011).
- [Mairesse 07] Mairesse, F., Walker, M. A., Mehl, M. R. & Moore, R. K. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. *J. Artificial Intelligence. Res. (JAIR)* 30, 457-500 (2007).
- [Tanaka 14] 田中 宏季, サクリアニ サクティ, グラム ニュービグ, 戸田 智基, 中村 哲, 物語発話からの自閉症スペクトラム障害児と定型発達児の語彙と韻律の特性分析, 日本音響学会春期大会 (2014). (to appear)
- [Hansen 95] H.M. Hansen, *Glottal characteristics of female speakers*, Harvard University, Ph.D. dissertation (1995).
- [Theodora 13] Chaspari, T., Gibson, D. B., Lee, C.C. & Narayanan, S. S. Using physiology and language cues for modeling verbal response latencies of children with ASD. *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3702-3706 (2013).