

## 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\*

☆ 田中 宏季, サクリアニ サクティ, グラム ニュービッグ, 戸田 智基, 中村 哲 (奈良先端大)

### 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, 1946年にレオカナーにより初めて報告された [1]. 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) こざわりと想像力の障害, の2つをあげており, 3歳から診断が可能であると言われている [2]. 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 の中心となる特性であると言われている. 社会性とコミュニケーションの障害の度合いを定量的に測定する事により, ASD の特性の理解, スクリーニングへの応用, また治療法・教育法の方角性策定が可能になると考えられる [3].

ASD と定型発達 (TD) 者の違いを明らかにする研究は数多く行われており, 社会性とコミュニケーションにおいては, ジェスチャーに関するもの [4], アイコンタクトや笑顔の頻度に関するもの [5], 韻律に関するもの [6], 声質に関するもの [7], 予期せぬ語彙に関するもの [8] などがあげられる. この中でも本研究では, 韻律と語彙に関して ASD の特異性について明らかにすることを目的とする.

ASD の韻律に関しては, レオカナーも報告している通り, モノトニック (単調) な, あるいはロボットの様な韻律になるという表現がこれまで使用されている [1, 9]. また語彙に関しては, Linguistic Inquiry and Word Count (LIWC) [10] を使用し, ASD と TD の成人におけるブログを解析した結果として, ASD 者では TD 者と比較して社交的な語彙を使用する頻度に個人差が多い事が報告されている [11]. 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. 個人の性格を表す内向性・外交性という尺度において, LIWC と韻律をまとめて特性分析と自動識別を行った報告 [12] があることから, 本研究ではその枠組みを ASD と TD との特性分析に応用する.

本研究では小児の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, 小児において物語発話 (ナラティブ) を使用する有効性が報告されている [13]. 本研究では小児における自らの印象に残った

体験の発話を題材に, ASD と TD の語彙と韻律について特性分析を行う.

### 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, ピ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 Table 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.

Table 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 Corpus [14] 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. 5分間における子どもの発話数は, S1からS6までそれぞれ, 30, 56, 32, 57, 59, 57である. ASD のグループ内では個人差が見られるが, ASD と TD において有意差は存在しない. 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.

\*Using Linguistic and Prosodic Cues for Understanding Nature of Autism Spectrum Disorders and Typical Development in Narrative. by Hiroki Tanaka, Sakriani Sakti, Graham Neubig, Tomoki Toda, Satoshi

### 3 特徴量抽出

本研究は先行研究 [12] に基づき発話毎に言語特徴量と音声特徴量をそれぞれ抽出した。抽出した特徴量は Table 2 にまとめている。さらに LIWC の辞書情報に加えて、1 発話中の笑いの頻度を新たな特徴量として加えている。

Table 2 語彙および韻律特徴量の情報

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, 知覚プロセス 生理プロセス, 関係について述べる割合
個人的関心	仕事, 達成, レジャー, 家庭について述べる割合
パラ言語情報	同意, 言いよどみ, フィラーの割合
基本周波数	標準偏差, 変動係数
パワー	標準偏差, 変動係数
発話速度	発話時間当たりの単語数

#### 3.1 言語特徴量

語彙を抽象化してカテゴリ化するためのツールである LIWC [10] を使用して言語特徴量を抽出した。日本語 LIWC は現在のところ開発されていない為、次のような手順で書き起こし文と LIWC の対応をとった。発話に対して MeCab<sup>1</sup> による形態素解析を行った後、形態素毎で日本語 Wordnet<sup>2</sup> を用いて検索をかけ英語に対応付けし、その結果に対して LIWC 辞書との対応をとった。また LIWC において英語のみ表れる情報および、サブカテゴリの情報においては特徴量として使用していない。

#### 3.2 音声特徴量

音声特徴量に関しては、個人差が表出される平均などの統計量は抽出せず、基本周波数とパワーにおける標準偏差 (fsd, psd) と変動係数 (fcov, pcov) の 2 つの尺度を抽出した。単語数を発話時間で割った、発話速度 (SR) も同様に抽出した。基本周波数とパワーの抽出に関しては、Snack Sound Toolkit<sup>3</sup> を使用した。

#### 3.3 射影正規化

ここで、特徴量を値レンジ [0, 1] に正規化を行った。0 は 1 話者中の最小値であり、1 は最大値である。i を発話とし、j を特徴量とした際、正規化前の値  $v_{ij}$  は、
$$p_{ij} = \frac{v_{ij} - \min_j}{\max_j - \min_j}$$
 により正規化される。

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

<sup>2</sup><http://www.omomimi.com/wjpn/>

<sup>3</sup><http://www.speech.kth.se/snack/>

### 4 特性分析

正規化後の特徴量に関して、t 検定、主成分分析、因子分析、決定木による特性分析を行った結果を報告する。統計処理には R<sup>4</sup> を使用した。

Table 3 は、各特徴量において t 検定による有意な差が見られた平均値の大きいグループを表している。これより、ASD は TD と比較して、6 文字以上の単語 (6 let.)、同意語 (assent: 「うん」等)、フィラー (filler: 「ええ」等) の使用が有意な差をもって多い事がわかる。さらに TD は ASD と比較して、社交語 (social: 「友だち」等)、感情語 (affect: 「楽しかった」等)、認知語 (cognitive: 「わかった」等) の使用が有意な差をもって多い、加えて f0 に関する標準偏差が大きいことも確認できる。その他の特徴量に関しては、p 値 0.01 以下での有意差は確認されなかった。

次に分散が大きく全体に寄与する特徴量を調査するため、主成分分析と因子分析を行った。主成分分析の結果として、第一主成分と第二主成分のバイプロットを Fig. 1 に示している。第二主成分までの全体への寄与率は 18% となっている。これより特に f0 に関するものと、パワーに関する成分で大きな分散をもっていることが確認できる。さらにフィラーに関しても第二主成分で大きな値となっている。主成分分析の軸を回転させた結果を得るため因子分析を行った。回転にはバリマックス法を使用した。因子分析の結果でも (Fig. 2 参照)、基本周波数とパワーが大きな値となったが、その他の言語特徴量 (WPS, 同意語など) も大きな値を示している事が確認できる。ここで、値の近い基本周波数の fsd と fcov および、パワーの psd と pcov の相関係数を調べると、それぞれ 80% ( $p < 0.01$ ) 以上の相関があった。これらの特徴量に関しては、次節の識別において標準偏差に関するもののみを使用した。

また決定木で IF-THEN ルールによる ASD と TD を分類するための重要な特徴量についても分析を行った。Fig. 3 に葉数を 10 とした際の結果を示す。これより全体としては韻律特徴量が多数をしめているが、語彙に関しては WPS がより分類に重要な特徴量だと確認できる。

### 5 ASD と TD の自動識別

ここで、ASD のスクリーニングに向けた自動識別に取り組む。前節での両グループの特性分析により、特徴量について次のセットを用意した: 1) 言語特徴量のみ、2) 音声特徴量のみ、3) 全ての特徴量、4) t 検定、

<sup>4</sup><http://www.r-project.org>

Table 3 それぞれの特徴量で有意差をもって大きい平均値のグループを表している. (\*:  $p < 0.01$ , \*\*:  $p < 0.005$ )

WPS	6 let.	laughter	adverb	pronoun	conjunctions	negations	quantifiers	numbers
-	ASD*	-	-	-	-	-	-	-
social	affect	cognitive	perceptual	biological	relativity	work	achievement	leisure
TD**	TD*	TD**	-	-	-	-	-	-
home	assent	nonfluent	fillers	fsd	fcov	psd	pcov	speech rate
-	ASD**	-	ASD*	TD**	TD*	-	-	-

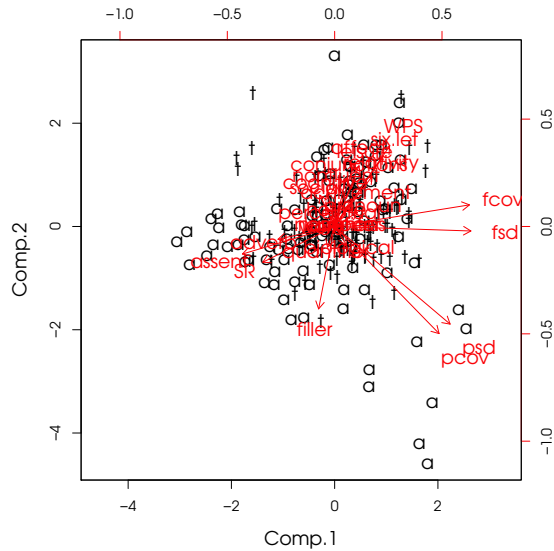


Fig. 1 主成分分析の結果. 第一と第二主成分を表示. (a: ASD, t: TD)

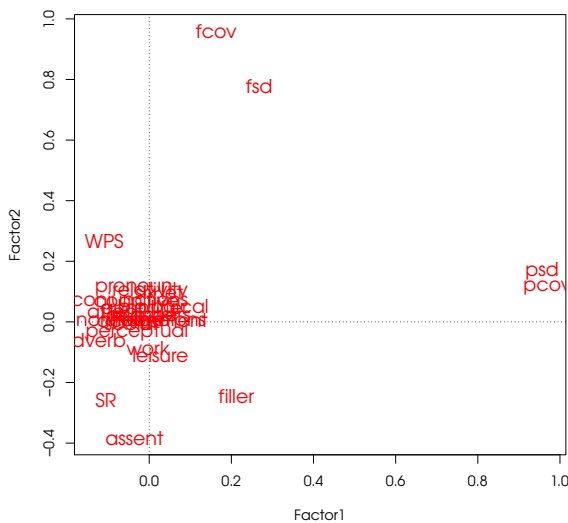


Fig. 2 バリマックス法を使用した因子分析の結果. 第一と第二因子を表示.

主成分分析, 因子分析, 決定木で有意差あるいは重要な特徴量として表れた特徴量 (Feature selection), 5) t 検定で有意差が出たもののうち, 互いに相関関係が見られない特徴量 (Feature selection (t-test + cor.)). Feature selection (t-test + cor.) の特徴量セットは, 6 let., social, affect, cognitive, fillers, assent, fsd である. ベースライン (BASE) をチャンスレートとした. Naive Bayes (NB) と線形カーネルの SVM を使用した交差検定 (cross=10) で識別を行い, さらに Leave-one-speaker-out 交差検定も行い, 正解率を算出した.

結果として, Feature selection (t-test + cor.) と SVM による分類が Leave-one-speaker-out において 66.7% で最も良い正解率を得た. しかしながら, 話者毎の正解率においては, S1 から S6 でそれぞれ, 78%, 60%, 53%, 51%, 82%, 78% となり, 個人差があることも確認された.

Table 4 Naive Bayes と SVM による ASD と TD の識別正解率. ベースラインと比較した際の有意差を表示している. (†:  $p < 0.1$ , \*:  $p < 0.01$ )

Feature set	Unweighted Accuracy [%]		
	BASE	NB	SVM
LIWC		62.2 †	70.3*
Prosody		57.6	67.6*
All	50.0	65.0 †	68.8*
Feature selection		67.4*	71.9*
Feature selection (t-test + cor.)		67.8 †	68.1 †
Leave-one-speaker-out	50.0	65.5 †	66.7 †

## 6 おわりに

本研究は, 物語発話による ASD と TD の特性の違いに関して分析を行った. 言語情報および韻律情報に関する特徴量を抽出し, その両方において ASD と TD での有意差が確認された. ASD では, 6 文字以上の難解な語彙を使用する傾向, 同意語やフィラーなどの使用の頻度が多くなる, またモノトニックなイントネーションになるという傾向が表れた. さらに社交語, 感情語, 認知語が比較的少ないことも明らか

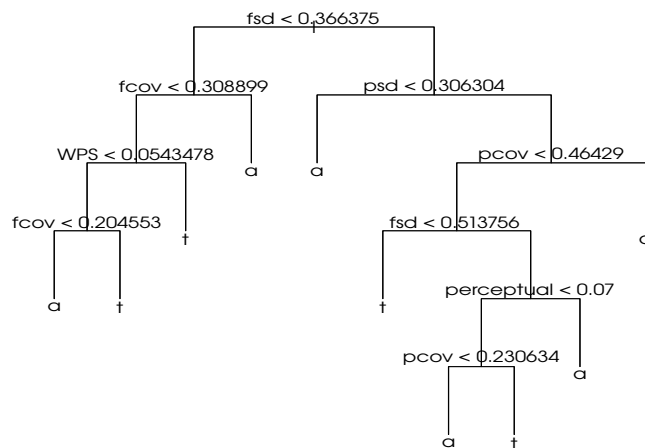


Fig. 3 葉数が10の決定木による重要な特徴量の選定. (a: ASD, t: TD)

となった。これらの特徴により、ASDの社会性とコミュニケーションにおける特異性が表出されているのだと予想することができる。これらの特徴を用いたASDとTDの識別では、6割以上の正解率を得る事ができた。しかしながら、大きく個人差が確認され、識別の確信度あるいは定量的な数値との関連を検討する必要がある。今後は、本パイロット実験を被験者を増やし発展させることに加え、母親の質問に対する応答や、予期しない単語の使用に関して特性を分析し、ASDのスクリーニングおよび治療の方向性策定に応用していく。

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る。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

[1] Kanner L. Autistic disturbances of affective contact. *Nervous Child* 2, 217-250 (1946).

[2] American Psychiatric Association. *The diagnostic and statistical manual of mental disorders*, 5. Washington, D.C. (2013).

[3] 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).

[4] de Marchena, A. & Eigsti, I. Conversational gestures in autism spectrum disorders: asynchrony but not decreased frequency. *Autism Research* 3, 311-322 (2010).

[5] Dawson, G., Hill, D., Spencer, A., Galpert, L. & Watson, L. Affective exchanges between young autistic children and their mothers. *Journal of Abnormal Child Psychology* 18, 335-345 (1990).

[6] McCann, J. & Peppe, S. Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders* (2003).

[7] Asgari, M., Bayestehtashk, A. & Shafran, I. Robust and Accurate Features for Detecting and Diagnosing Autism Spectrum Disorders. (2013).

[8] 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).

[9] Paul, R., Augustyn, A., Klin, A. & Volkmar, F. R. Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 35, 205-220 (2005).

[10] Pennebaker, J. W., Francis, M. E. & Booth, R. J. *Linguistic inquiry and word count: LIWC [Computer software]*. Austin, TX: liwc.net (2007).

[11] Newton, A. T., Kramer, A. D. I. & McIntosh, D. N. Autism online: a comparison of word usage in bloggers with and without autism spectrum disorders. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 463-466 (2009).

[12] 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. Artif. Intell. Res.(JAIR)* 30, 457-500 (2007).

[13] Davis, M., Dautenhahn, K., Nehaniv, C. L. & Powell, S. D. Towards an Interactive System Facilitating Therapeutic Narrative Elicitation in Autism. (2004).

[14] Mower, E., Black, M. P., Flores, E., Williams, M. & Narayanan, S. Rachel: Design of an emotionally targeted interactive agent for children with autism. *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, 1-6 (2011).