

幼児の簡易語彙能力チェックリスト作成における 幼児分類の効率化

塚田 元春^{*1} 南 泰浩^{*2} 小林 哲生^{*3} 奥村 優子^{*4}

^{*1*2} 電気通信大学大学院 情報システム学研究所

^{*3*4} NTT コミュニケーション科学基礎研究所

^{*1*2} {tsukada, minami.yasuhiro}@sd.is.uec.ac.jp

^{*3*4} {kobayashi.tessei, okumura.yuko}@lab.ntt.co.jp

1 はじめに

我々は幼児の語彙発達を調べる手法として、2688項目の電子的語彙チェックリストをタブレット PC 上のアプリとして実装した [1]。このアプリを用いて、幼児の語彙能力の指標として、幼児の日齢を計算する。この日齢は、幼児の実際の日齢を表したものではなく、幼児の語彙能力に等しい語彙能力を持つ幼児の平均日齢を表し、その幼児の客観的な語彙能力を示す指標である。しかし、このチェックリストでは、母親が2688項目の回答をする必要がある。このため、項目数を削減しても幼児の語彙数を正確に求める手法が必要である。これまで、幼児の語彙獲得のタイプによって語彙数等に差があることが知られており [2, 3, 4]、幼児の語彙数の推定値をタイプによって補正することも必要であることが分かっている。

このようなチェックリスト語彙の削減手法として、我々は、簡易語彙チェックリストと幼児の語彙獲得タイプを分類して補正を行う手法を提案した [5]。しかし、この手法では、40項目の簡易語彙チェックリストに対し、補正のための幼児の語彙獲得タイプの分類に400項目が必要であった [5]。回答者の負担を減らすためには、さらに、この分類に必要な項目の削減が必要である。

本研究では、Q3 統計量を利用する語のクラスタリングによる幼児の語彙獲得タイプの分類を効率的に行う手法を提案する。

2 データセットと検証方法

手法の効果を確認するため、日本語 2688 語の項目からなる NTT 語彙チェックリスト [1] を用いて、1040 人の幼児から得られた結果を利用した。本稿では、1040

人のうち 50 人分の回答データを能力値である日齢の推定用として利用した。残りの 990 人分を語彙チェックリスト作成のために利用する。データの信頼性を高めるため、重複しない 50 人分の回答データを 20 回繰り返して取り出し、その他のデータを語彙チェックリスト作成に使用して、その結果を検証する。

また、推定日齢と実際の日齢の差の分散をもってチェックリストの推定精度を評価する。20 回の検証の分散の平均を比較する。この値が小さいほど精度は良い。

3 従来手法

従来手法の補正方法 [5] は、簡易語彙チェックリストによるクラスタ毎で幼児に対して推定日齢 \bar{x}_k と実際の日齢 d_k の平均値を計算し、その差を補正項にする方法である。

3.1 補正方法

まず、語彙日齢が若い W 語の品詞チェックリストを用意する。各項目には社会語、名詞、述語、閉じた語、その他というカテゴリ [6] が振られている。幼児 k のカテゴリ毎の発話語彙数をそれぞれ $V_{k,s}, V_{k,n}, V_{k,p}, V_{k,c}, V_{k,o}$ とする。また、 N 語の品詞チェックリストより得られる幼児 k の発話語彙数を V_k とし、カテゴリ毎の発話割合を 5 次元のベクトル

$$v_k = \left(\frac{V_{k,n}}{V_k}, \frac{V_{k,s}}{V_k}, \frac{V_{k,p}}{V_k}, \frac{V_{k,c}}{V_k}, \frac{V_{k,o}}{V_k} \right) \quad (1)$$

で表す。

幼児を (1)12ヶ月以下, (2)13-18ヶ月, (3)19-24ヶ月, (4)25-30ヶ月, (5)31-36ヶ月, (6)37-48ヶ月の月齢グ

ループに分ける。月齢グループ毎に、ベクトル v_k を k-means 法で二つのクラスに分類する。クラス毎に幼児の推定日齢と実際の日齢の差 $\bar{x}_k - d_k$ の平均を計算し、この値をクラス毎の補正值とする。

3.2 補正効果の検証

ここでは、実際の日齢と推定日齢間の分散を計算し、その値が小さくなることによって、性能の良さを評価する。これは、能力推定が上手くいっていれば、実際の日齢の周りに小さく推定日齢が分布することになるであろうという予測による。補正をせずに、40 語の簡易語彙チェックリストだけを用いた場合の平均分散は 17832 となった。語彙日齢が若い W 語の品詞チェックリストによって補正した結果を表 1 に示す。分散の最小値は 900 語の場合の 9874 である。

表 1: 従来手法の品詞チェックリスト語数と平均分散

| 語数 W | 平均分散 | 語数 W | 平均分散 |
|--------|-------------|--------|-------------|
| 50 | 11218 | 600 | 9953 |
| 100 | 10426 | 700 | 10011 |
| 200 | 10405 | 800 | 9973 |
| 300 | 10611 | 900 | 9874 |
| 400 | 10016 | 1000 | 9899 |
| 500 | 9981 | | |

4 提案手法

従来手法の品詞チェックリスト内には、幼児の能力推定の観点から同じ役割を持つような単語が存在する可能性がある。本稿では、品詞チェックリスト内の単語を Q_3 統計量 [7] によりクラスタリングし、各クラスから語を選出することで、品詞チェックリストの語数を減らす手法を提案する。

4.1 Q_3 統計量

Q_3 統計量とは、テストなどで、二つの問題（項目）が同時に理解されやすいかどうかを示す指標であり、任意の 2 項目間において、幼児の得点とその項目の難易度の残差の相関として定義される指標である。本稿の場合、幼児の得点と難易度の残差とは、幼児が実際にその単語が理解できたかという 1 か 0 かの数字から

簡易語彙チェックリストによって推定されたその単語の平均正答率の差であり、難易度の残差が大きいほど、その単語に対してどれだけその幼児が他の子どもに比べてよくできるかを表している。平均正答率を引き去ることにより、日齢による影響を取り除くことができるため、 Q_3 統計量は同時期に習得される単語を日齢の影響を受けずに観測できる。

4.2 補正方法

まず、単語をクラスタリングするために、全単語間の Q_3 統計量を計算し、これをグラフ化する。各単語を頂点とし、単語間を辺とする。 Q_3 統計量が辺の値となる。 Q_3 統計量を用いた単語のクラスタリングは、結びつきの弱い部分でグラフを切断し k 個に分ける問題に帰着する。本稿では、spectral-clustering[8] を用いてこれを行った。

次に、クラスター毎に代表として若い n 語を取り出すことにより、品詞チェックリストを削減する。その後の補正については従来手法と同様の手法を用いる。

4.3 補正効果の検証

40 語の簡易語彙チェックリストは従来手法の検証と同じものを使う。この平均分散は 17832 である。また、クラス数 k を 20, 40, 60, 取り出し数 n を 1, 2, 3, 4 とし、これらを組み合わせて性能の検証を行った。語彙日齢が若い W 語の品詞チェックリストとして、事前実験で最良の結果を出した $W = 600$ のチェックリストを削減前のものとした。補正した結果を表 2 に示す。語数が大幅に削減されているにも関わらず、多くの条件で従来の最良値である 9874 を上回っている。クラス数について、 $k = 60$ の結果が悪いことから、 $k = 40$ 程度で良いことが考えられる。

5 おわりに

本稿では、幼児の語彙獲得タイプの分類チェックリストを、 Q_3 統計量を利用した語のクラスタリングによって削減し、より効率的な幼児の語彙能力チェックリストを提案した。

さらに、分類チェックリストの語数を削減した場合でも、推定精度が向上することが確かめられた。今後は、幼児分類において、重要となる語について、より詳細に調査したい。

表 2: 提案手法の条件と平均分散

| 条件 | 語数 | 平均分散 |
|-----------------|-----|-------------|
| $k = 20, n = 1$ | 20 | 9364 |
| $k = 20, n = 2$ | 40 | 9443 |
| $k = 20, n = 3$ | 60 | 9024 |
| $k = 20, n = 4$ | 80 | 9169 |
| $k = 40, n = 1$ | 40 | 9222 |
| $k = 40, n = 2$ | 80 | 9266 |
| $k = 40, n = 3$ | 120 | 9273 |
| $k = 40, n = 4$ | 160 | 9556 |
| $k = 60, n = 1$ | 60 | 9971 |
| $k = 60, n = 2$ | 120 | 9505 |
| $k = 60, n = 3$ | 180 | 9938 |
| $k = 60, n = 4$ | 240 | 9882 |

in English and Italian,” *Journal of Child Language*, vol. 26, no. 1, 69-111, 1999.

- [7] W.M. Yen, “Effects of local item dependence on the fit and equating performance of the three-parameter logistic model,” *Applied Psychological Measurement*, vol.8, no.2, 125-145, 1984.
- [8] Jianbo Shi and Jitendra Malik, “Normalized Cuts and Image Segmentation,” *IEEE Transactions on PAMI*, Vol. 22, No. 8, 888-905, 2000.

謝辞

本研究の一部は、科学研究費助成事業基盤研究（B）17H02190 による援助を受けた。

参考文献

- [1] 小林哲生, 奥村優子, 南泰浩, “語彙チェックリストアプリによる幼児語彙発達データ収集の試み,” 電子情報通信学会技術研究報告 HCS, HCS2015-59, 2016.
- [2] 南泰浩, 小林哲生, 杉山弘晃, “語彙爆発の新しい視点: 日本語学習児の初期語彙発達に関する縦断データ解析,” *ベビーサイエンス*, 12, 34-49, 2013.
- [3] 南泰浩, 小林哲生, 杉山晃弘, “語の学習では本当に幼児は名詞を早く習得する? 一語の理解・発話日齢の推定による名詞優位性の言語間比較一,” *赤ちゃん学会第 13 回学術集会*, P-30, 2013.
- [4] 南泰浩, 小林哲生, “幼児早期出現語の理解-発話指標による名詞学習の優位性の検証,” *言語処理学会第 19 回年次大会*, 548-551, 2013.
- [5] 森山佑介, 南泰浩, 小林哲生, “幼児の能力推定のための簡易語彙チェックリストの提案,” 電子情報通信学会技術研究報告, 145-150, 2017
- [6] C. Caselli, P. Casadio, E. Bates, “A Comparison of the Transition from First Words to Grammar