

親子共学の間からの発話に着目した幼児学習コーパスの構築

Construction of an infant's learning corpus based on the observation of utterance in place of parents and children coeducation

石川 翔吾^{*1} 桐山 伸也^{*2} 北澤 茂良^{*2}
Shogo Ishikawa Shinya Kiriyaama Shigeyoshi Kitazawa

^{*1} 静岡大学大学院情報学研究科
Shizuoka University, Graduate School of Informatics

^{*2} 静岡大学情報学部
Shizuoka University, Faculty of Informatics

In this research, we observed the infant's behavior in the infant classroom, to improve the quality of the infantile education that followed the idea of the parents and children coeducation. We constructed an infant's learning corpus, and succeeded in obtaining valuable findings for an educational improvement, by analyzing and describing the reactions of children in the classroom, expressed by infant's utterances.

1. はじめに

幼児期において子供の触れるものは、大抵の場合親に委ねられている。すなわち、幼児の健やかな成長を促進するためには、親は子育てに対する知識やノウハウが必要である。この観点から、幼児教育には子供だけの教育にあらず、その親もまた教育に関する知識を得ることが重要視されている。また、愛着による理論[1990 M.Minsky]によると、愛着によって学習が進むような特定の学習方法があり、幼児がアイデンティティを形成するのに最もお手本となるモデルが大抵の場合親になることから、共学的重要性が言える。

そこで、親子の学習過程を収録し、幼児教室における幼児の発話を主な手がかりに教育に対する反応を観察・分析・記述していくことで、幼児学習コーパスを構築する。コーパスは発話を伴う行動、波形データ、書き起こし、韻律情報、発話の意図などのトラックを独立に設けた幅広く、柔軟性を持った仕組みを持つ。

さらに、幼児発話の特徴を分析し、本研究でメインとなる観察の支援をする。発話を検索の対象とし、アノテーションしたいシーンを抽出するようなシステムを構築する。現在その第一段階を実装することができた。

本研究では、幼児学習コーパスで解析した幼児の行動を下に、幼児教育へフィードバックできる知見を得ることを目的とする。

2. 発話の分析

2.1 幼児発話の意図の解釈

アノテーション作業には、膨大なデータ量にアノテートしていくため、効率化が必要となる。そこで、発話を検索の対象としたアノテーション支援システムを作成する。

発話の分析には、特に幼児の非言語的発話に着目してその音声を持っている音響的特徴を分析する。幼児は語彙力が少ないため、非言語的発話が圧倒的に多い。そのため、非言語的発話には、大人とは違い発話に意図が含まれている可能性が高い。そこで、遊びや道具を使いながら発した言葉を収集し、意味を解析しコーパス化する。現在 200 程度の発話を抽出し、その意味を解析した結果、非言語的発話を状況で使い分ける傾向があり、特に母音の「あ」、「お」、「ん」の使用頻度が高く、その非言語的発話を状況に応じて、強弱、高さ、つまり韻律を

連絡先: 石川 翔吾, 静岡大学大学院情報学研究科,
cs1007@s.cs.inf.shizuoka.ac.jp

使い分ける傾向があることが分かった。

2.2 音声区間検出実験

アノテーション支援システムを第一段階となる、音声区間の検出を行った。これにより、幼児発話の意図の解釈をするための効率を図ることが可能になる。

そこで、パワーをパラメータとした音声区間検出プログラムにより約 60 分のデータの音声区間の検出実験を行い、精度を検証した。正解ラベルとして、予めつけておいたラベルと比較することによって、評価する。まず、正解ラベルの基準を以下に示す。

- ・ ラベルは、音声区間とポーズとに区別
- ・ ポーズの区間は 500ms 以下とする
- ・ 雑音もポーズとする

これらの基準を下に作成した正解ラベルと、自動検出したラベルとを比較する。一部の例を図1に示す。ラベル層の一番上が正解ラベルとなり、その下から閾値を変えた自動検出ラベルとなる。下のラベルほど閾値が大きい。

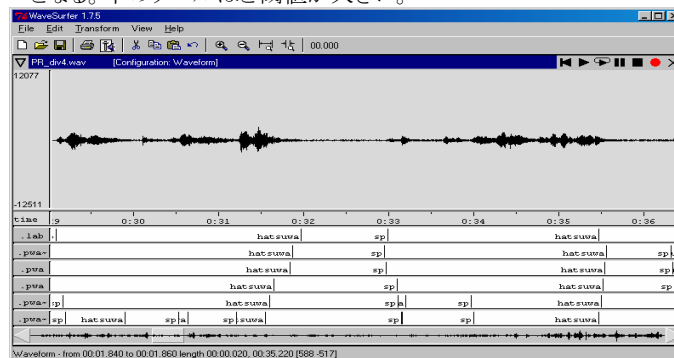


図 1: 各閾値の音声区間検出

ここでは、ずれ幅を設けず、正解ラベルと少しでも重なりと正解とする。今回の音声区間検出の目的は音声だけをいかにうまく検出し、アノテーションを支援するための前処理を施すことである。そのため、音声区間を含むラベルを検出すると、例えばずれ幅が大きかったとしても、正解とする必要がある。このような正解の基準を用いると、例えば、ある音声区間を2つに分けて検出した場合などは、正解が2つとして算出されるため、精度の評価には適合率を用いて算出する。図 2 に閾値を 12~30 まで変化させたときの閾値と正解精度の関係図を示す。適切な音声区間、ポーズを得ることが重要なため、精度の指標として正解精度を算出する。正解精度は、以下の式で表される。

$$\text{正解精度} = \frac{\text{正解数} + (\text{ポーズ数} - \text{誤検出数})}{\text{全ラベル数}}$$

ここで、誤検出とは、正解の音声区間をポーズと検出した部分のことを指す。さらに、300ms 以上正解の音声区間に重なった場合、または途中の音声区間に表れる場合とする。

閾値を小さくすると正解精度は良くなるが、ポーズ区間を検出することができず、音声区間検出という目的とは言えない結果になる。閾値が大きいとその逆のことが言える。結果、閾値を18に設定することが適切であることが分かる。

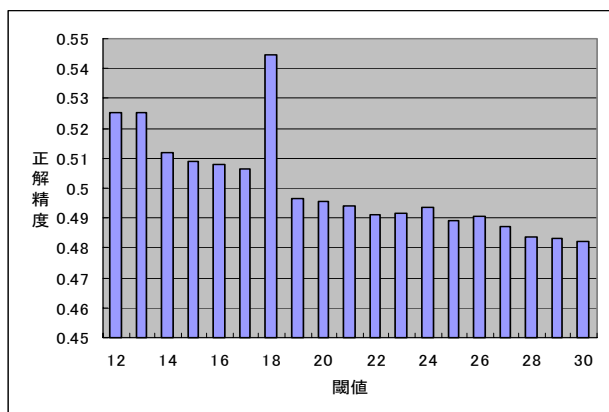


図2: 閾値ごと(12~30)の正解精度の関係図

3. 幼児行動の記述

幼児の発話を主として、行動を自然言語で記述・解析していくことで、コーパス化する。アノテーション作業には、DVD オーサリングシステム(図3参照)を用いて記述する。このシステムを用いることによって、教育者、親、専門化などの幅広い立場の人が記述することができ、コーパスに蓄えられた知識を共有することができる。



図3: オーサリングシステム

アノテーションの方針として、人工知能の第一人者であるMinskyの提唱するSix-Level model[forthcoming M.Minsky]に基づいて、幼児の行動の裏側に隠れている心的な要素を記述していく教材を使っているときの幼児の発話に表れる反応を、このモデルに則して記述していく。(図4参照)。このアノテーション結果により、幼児の教育コンテンツや道具などに対する反応を解析し、コーパスへ反映させた。

ここで、親や専門家コメントというのが重要となる。コメントを下に心的状態の記述の裏づけを取ることによって、より正確な心的要素を記述していくことができる。現在、親や専門家が自由にアノテーションできる環境も整ってきている。[2005 川口]

R君は「大丈夫?もう治った?」と言って、機嫌を伺う。		
Six-Level model	tag	心理状態
self-conscious thinking	社会的	人を泣かせてはいけない。 仲良く遊ばないといけない。
self-reflective thinking	自制	自分本位ではなく、 他人を意識して行動しよう。
deliberative thinking	推測	怒って誰かに言わないか?
learned reactions	経験	後で誰かに怒られたことがある。
親のコメント		
保育園・家庭での教育: 仲良く遊ばなければいけない。		
保育園・家庭での経験: 仲直りしたと思っていても告げ口されると怒られるかもしれない。		

図4: アノテーション例

4. コーパスの分析

コーパスの分析にはまず、自然言語処理によりコーパスに記述した概念を抽出する。[2005 大竹]その概念を下に評価するための手掛かりとし、さらに発話と結びつけることによって、コーパスに記述された知識を分析していく。そうすることにより、特定の概念を検索するだけで、知見を探し出すシーンを絞ることが容易になる。以下に例を示す。

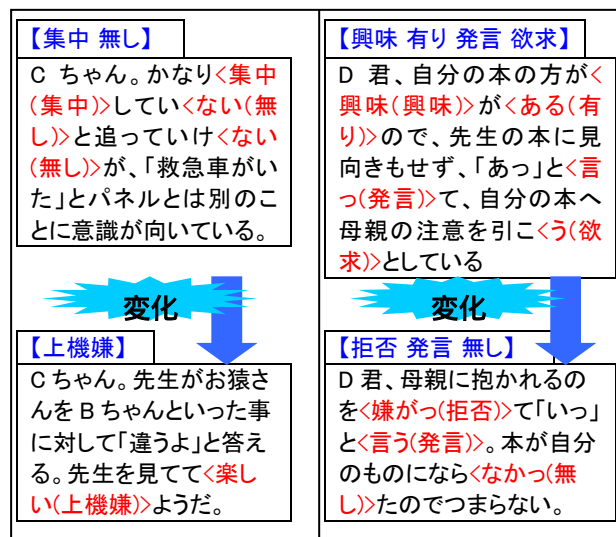


図5: 自然言語処理を用いたコーパスの評価

ここで、途中の変化というものが大事になる。上図の例の2つの変化の状況を観察すると、以下のようになる。

変化①:

自信がないが、答えた回答が合っていて褒められた。次の回答から大きな声で答える。

変化①では、幼児の学習に対する心境の変化が大きく見られる。答えることができ、褒められたことから自信となり、結果的に楽しく学習することに繋がっている。

変化②:

本をしばらく独占しているが、先生に読んでもらうため本を渡す。そのため、他の子がその本に興味を持ち所有権争い。

変化②では、独占欲の強い子が、所有権が移ってからの態度の変化が観察される。つまり、まだ本に執着がある場合は例のように不満になるが、ない場合は飽きとして観察される。

このように、発話に関するアノテーションから自然言語処理によって見当を付け、行動に変化が表れる部分に着目して観察す

ることにより、変化①では教育の仕方への知見、変化②では教材に対する知見というものを得ることができたと言える。

5. まとめ

親子共学の場合を収録し、主に発話情報から幼児の反応をアノテートし、教育の質の改善を目指した「幼児学習コーパス」を構築し、自然言語処理を施すことで効率的に幼児教育にフィードバックできる知見を得ることができた。

今後は、定期的に幼児教育を実施していき、音声、自然言語、しぐさなどのトラックを上手く融合させることによって知識を溜め、抽出する枠組みを作っていく、根源的コモンセンス解明への足がかりとしていく。

参考文献

- [1990 M.Minsky] M.Minsky, 安西祐一郎訳: ‘心の社会’, 産業図書 (1990)
- [forthcoming M.Minsky] M.Minsky: ‘The Emotion Machine’, (forthcoming)
- [2005 大竹] 大竹康太, 川口紗季, 石川翔吾, 大谷尚史, 山本剛, 坂根裕, 桐山伸也, 竹林洋一: ‘親子共学のための自然言語処理によるマルチモーダル知識収集支援’, 第4回情報科学技術フォーラム, (発表予定)
- [2005 川口] 川口紗季, 坂根信一, 坂根裕, 竹林: ‘親子のコミュニケーションを促進する幼児教育の一検討’, 第67回情報処理学会全国大会, (2005)