

マルチモーダル行動コーパスを用いた音声言語獲得過程の分析 Analysis of language acquisition process using Multimodal Infant Behavior Corpus

川崎 壮太*1 桐山 伸也*2 竹林 洋一*2 北澤 茂良*2
Souta Kawasaki Shinya Kiriyama Yôuchi Takebayashi Shigeyoshi Kitazawa

*1 静岡大学大学院情報学研究科 *2 静岡大学情報学部
Shizuoka University, Graduate School of informatics Shizuoka University, Faculty of Informatics

This paper describes a multimodal infant behavior corpus by analyzing language acquisition process. Based on Analysis of special Syllables, three factors were analyzed. These factors consist of utterances such as, “yarutte”, “natikettearuno” and “nanikaattaranandesuka”. These results provide useful findings with phonetic perception and construct of social nature model. The results of our analysis show that a multimodal infant behavior corpus is useful in analyzing language acquisition process.

1. はじめに

人間がコミュニケーションをとる際に用いられる音声言語は重要である。もちろん表情や目線、動作などもコミュニケーションをする上では必要であるが、相手とコミュニケーションをとろうとする際に、乳幼児から大人まで共通して用いられる多くの要素は音声言語である。

人間は怒ったり、質問したりする際にイントネーションを変えたり、単語を取捨選択することにより、シチュエーションに合わせた音声言語を活用する。音声言語はその名の通り、音声を持つ要素と言語を持つ要素によって構成がなされており、人間はその両方の要素をうまく用いてコミュニケーションをとっている。

近年、この人間の音声言語に関する研究が盛んに行われている。その中でも幼児を対象とした研究も多く存在し、幼児の音声言語獲得に関する知見も増えている。しかしそれらの多くは実験的に集められたデータを対象としたものが多い。また実験的ではなく、幼児の自由な発話を対象とした分析も進んでいるが[麦谷 2004]、音声言語獲得過程の詳細な結果は得られていない。

幼児の音声知覚の知見を得る方法として選考振り向き法から、幼児が特定の音素配列に興味を示す事がわかっている[佐藤 2007]。しかし幼児がどのようにその知覚した音声知識を持って音声言語を産出しているかについては詳しい研究がなされていない。

本稿では音声の産出の詳細分析を視野に入れながら、佐藤らの知覚に関する検証から、特殊音に焦点を当てた発話分析を行う。多様なインタラクションにおける発話分析を基に、複数の分析事例を挙げ、マルチモーダル行動コーパスが音声言語獲得過程の分析に有用であることを示す。

2. 音声言語獲得過程分析の背景

2.1 音声言語獲得過程の分析方針

人間の音声言語は音素、パワー、ピッチから構成されている。また人間が話す言葉は1文字とは限らないので、音素の組み合わせや文法なども含めて人間は音声言語を認識している。

本稿で対象としている幼児も母親など大人の発話を知覚して音声言語を獲得している。しかし幼児は大人と違い、単語や文

法といった知識をもっていない。このことから幼児を対象とした音声言語獲得過程を分析することにより、幼児が音声という多様な情報を持つ物のどのような部分の特徴を捉えて音声言語を獲得しているのか(ディステインクティブフィーチャー)を解明できると考える。

本稿では幼児の音声言語獲得過程において、幼児のディステインクティブフィーチャーの解明を目標に、マルチモーダル行動コーパスを用いて、他者との多様なインタラクションからの発話分析を行う。

2.2 データ収録環境とその特徴

我々はマルチモーダル行動コーパスを構築している[桐山 2008]。このマルチモーダル行動コーパスには幼児の自然な環境下での動作と発話が記録されており、先生、親子、子供同士の多様なシチュエーションにおける発話が記録されている。また幼児が発話するまでの状況や発話時の環境を用いて、なぜ幼児がそのような発話をしたのか、幼児の発話状況を考慮した詳細な分析が可能であるという特徴も兼ね備えている。

個々に発話を分析するだけではなく、1人の幼児の経年的な観察が可能のため、音声言語獲得過程はもちろんのこと、社会的な行動発達過程など様々な知識獲得過程を調べられるという特徴も持っている。

2.3 使用分析ツール環境

我々は図1に示すアノテーションツールを用いることにより、幼児教室の映像を見ながら、幼児の発話ごとにアノテーションすることが可能である。また図2に示すように複数ある発話シーンを並べて表示し、クリック1つでその発話の波形を表示させる機能を持たせている。これにより時系列に発話を並べて観察、同じシチュエーションのもので並べて観察することが容易に可能である。またアノテーションした発話を音素別に検索し、並べすることも可能である。



図 1:アノテーションソフト

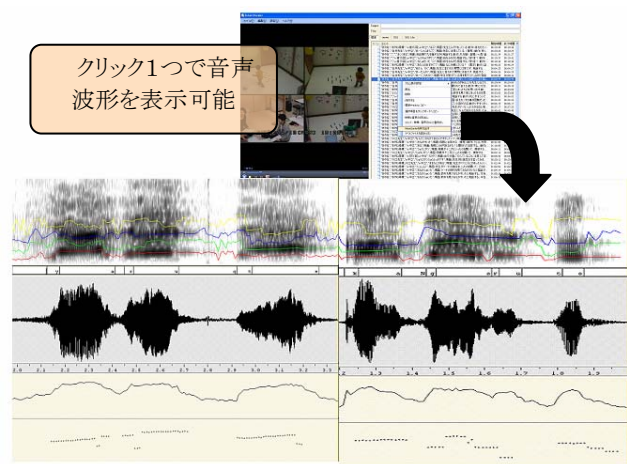


図 2:クリック1つで波形を表示可能

3. 音声言語獲得過程の分析

3.1 対象幼児と対象シチュエーション

分析対象とした幼児は 2 歳～5 歳になる幼児一人である。幼児が母親や先生と会話が多く発声しやすい教材(動物や物のイラストが入っている, または母親が幼児に発話を促すような教材)を使用している授業の映像を重点的に分析し, 現在 89 場面 400 発話の書き起こしデータを蓄積している。

特に着目したのは他者とのインタラクションからでた発話である。違う立場の人とのコミュニケーションの中で幼児は相手の言葉を聴き, 返答する。他者とのインタラクションの中で幼児は, 相手の音声言語を知覚し, それに対する音声言語を産出している。しかし幼児が大人の言葉をすべて知覚できているとは限らない, 幼児の中である知識をもって取捨選択を行っているはずである。筆者らはこの取捨選択のメカニズムを解明することにより, 幼児のディスティンクティブフィーチャーを見つける。

3.2 分析着眼点

1 章でも述べたように佐藤らの検証により, 第1音素の音圧と第2音素の音圧の差が大きい発話に幼児は興味を抱くということが検証されているので, 幼児にとって知覚しやすい音は発話しやすい音であるという知覚と産出の仮説を基に, 本研究では音圧差が大きい A+U という発話(第1音素が A を含む音で第2音素が U を含む音である発話)を筆頭に, 特殊音(撥音, 促音, 長音)にまで視野を広め, それに関わる発話の分析を中心に行う。

またディスティンクティブフィーチャーを見つけるために, 幼児が音声言語を知覚し, 産出する過程に関わる他者とのインタラクションの中で発話されたものを中心に分析する。

4. 分析結果

3章で述べたように特殊音, 他者とのインタラクションを基にした複数の発話の分析結果を以下に示す。

1つ目は「やるって」という発話である。発話状況は, 他の幼児が席をはずし, その幼児に向かって次の教材を実施するというのを伝える場面である。発話順序は幼児が他の幼児に向かって「やるって」と発話した後に先生が「やるって」と繰り返して発話し, それを聞いた幼児がもう一度「やるって」と発話する。幼児と他の幼児と先生とのインタラクションがある。1回目と2回目の発話に違いがあり, 図3に示すように「やるって」と正しく発話できていたのが「やるて」と促音を省略して発話している。これらの発話から, 幼児が先生とのインタラクションにより, 先生の発話のアクセントを知覚し, 促音を省略して発話をしていると考察できる。これにより幼児はアクセントに着目する傾向があるという仮説が立てられる。

2つ目は「なーちけてあるの?」という発話である。発話状況は手元にあったクレヨンにセロハンテープが貼ってあり, それに疑問を持った幼児が母親と先生に理由を質問する場面である。発話順序は幼児が母親に「なーちけてあるの?」と聞き, それに対して母親が答えるが, 幼児が納得できず, 次に先生に対して「なんでつけてあるんですかっていう」と発話する。幼児と母親, 幼児と先生のインタラクションがある。この発話の特徴は発話相手によって発話内容が変わるという点である。これらの発話から幼児の社会性モデルの構築に有用である。

3つ目は「なにかあったらなんですか?」という発話である。発話状況は, 買い物ごっこをする教材を使用しているときに, 先生の発話を模倣する場面である。発話順序は先生が幼児に向かって「なにかあったらいいですかきてみて」と発話し, それを聞いて「なにかあったらなんですか」と発話する。幼児と先生と母親のインタラクションがある。この発話は先生の発話の中で, 頭と終わりだけをつないで模倣するバスタブエフェクトの構造を持つ。

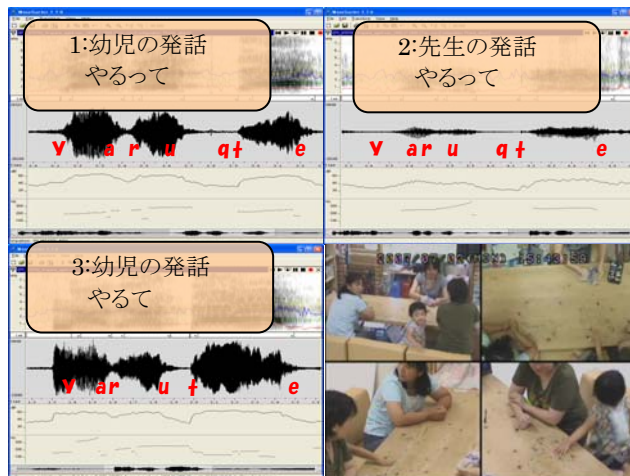


図 3:「やるって」という発話の分析結果

5. まとめ

幼児の音声言語獲得過程の分析を分析する際に, マルチモーダル音声行動コーパスを用いることによって, 自由な環境下での発話や多様なインタラクションからの発話が分析可能であ

る。多様なインタラクションという特徴を用いて、幼児の音声言語獲得過程におけるディスティンクティブフィーチャーを導き出すことを目的とした。佐藤らが検証した音圧差という知覚の観点から、筆者らは特殊音という観点で分析を行った。

実際にマルチモーダル行動コーパスから得られたデータを分析し、「やるって」「なーちけてあるの」「なにかあったらなんですか」という発話を分析した。これらの分析結果から幼児の知覚に関する知見、社会性モデルの構築、バスタブエフェクトに有用であることを導き出し、マルチモーダル行動コーパスが音声言語獲得過程の分析に有用であることを示した。

今後このマルチモーダル行動コーパスを活用して、母親の発話特徴量を検出し、母親の発話と幼児の言語獲得の相関関係という観点からも音声言語獲得過程の分析を進める。

謝辞

本研究は、総務省・戦略的情報通信研究開発推進制度(SCOPE)地域 ICT 振興型研究開発「マルチモーダル幼児教室を機軸とする成長する育児支援コンテンツとヒューマンネットワークの実現」の支援を受けて行った。

参考文献

- [表谷 2004]表谷綾子: 乳幼児の音声言語獲得, Technical report of IEICE. Thought and language , 104(316) pp.13-18 20040910
- [佐藤 2004]佐藤久美子, 梶川祥世, 坂本清恵, 松本博文: 日本語母語乳児の文中からの単語切り出しにおけるアクセントと音素配列の役割, Journal of the Phonetic Society of Japan, Vol.11 No.3 pp.38-47 December 2007
- [桐山 2008] 桐山伸也, 石川翔吾, 大谷尚史, 北澤茂良, 竹林洋一: 音声コミュニケーション研究のためのマルチモーダル音声行動コーパス, 情報処理学会研究報告, 72 pp.111-116 20080718.