

子どもの言語表現の発達の検討  
ー実験場面についての報告のテキストマイニングによる分析ー

The Developmental Study of Children's Verbal Reports :  
Analysis of Accounts about Experimental Situation by Using Text Mining

吉村 啓子\*

Keiko YOSHIMURA

抄 録

本研究の目的は、直前の経験に関する3歳から9歳の子どもの音声言語報告を、テキストマイニングの手法によって分析する事により、言語発達研究の一つの可能性を示すことにある。182名の子ども（年少児40名、年中児40名、年長児40名、小学校1年生24名、小学校3年生38名）が、簡単な記憶の実験を経験し、その場面について別室の実験者にその内容について報告するという場面を設定した。子どもの発話データを文字化し、テキストマイニングにより頻出語を抽出し分析を行った。その結果、「形」、「色」、「並べる」、「を」という言葉の使用に発達的な変化が見られた。また、小学校3年生になると「覚える」という言葉を使うようになることも明らかになった。

I 目的

本研究の目的は、直前の経験に関する3歳から9歳児の音声言語報告を、テキストマイニングの手法によって分析する事により、言語発達研究の一つの可能性を示すことにある。

子どもが自分の経験を報告する形のデータを扱った研究はあまり多いとは言えない。藤崎（1982）<sup>1)</sup>は、3歳から6歳までの子どもについて、報告場面における計画的構成について発達の検討を行っている。この研究は、休日の翌日に実験者が園に赴き、子どもが保育室で他の子どもたちと保育者の前で自分の経験を語る様子を記録し、分析したものである。また、尾山・仲（2013）<sup>2)</sup>では、5歳から6歳までの子どもが、ポジティブな出来事とネガティブな出来事についての問いに対して実験者に答え、それを記録し分析している。これらの研究から一定の知見は得られているものの、問題点としては、子どもの経験を統制できていない事があげられる。仲野・長崎（2012）<sup>3)</sup>によれば、自己の経験の報告についての研究が少ない理由として、「“お話作り”として観察場面を設定しやすいフィクションナラティブに比べ、自己の経験に関する自然な語りを引き出すのが困難なこと」を挙げ

---

\* 関西国際大学教育学部 教育総合研究所学内研究員

ており、「観察場面については十分に考慮しなければならない。」と述べている。そして経験を統制した上で、その経験を母親（別室にいて、子どもの経験は見えていない）に報告するという実験を試みている。経験とは、実験者と幼児が一緒にケーキを作り、実験者が卵を落とすなどのハプニングを起こすものである。子どもの語りを全て節に分割し、結束性という観点で認定基準を設けて分析している。その結果、4歳頃に時系列性が意識されるようになること、5歳頃には主観的視点からの意味づけがみられることが明らかにされた。

これらの研究から、子どもの経験の報告についての研究が少ないのは、経験を統制しにくいことに加え、子どもの言語報告を数量化して分析する方法が煩雑であることも大きな理由であるように思われる。さらに、研究者が先に決めた基準以上の分析はできず、多くの価値あるデータを分析対象から外している可能性が考えられる。

大人の文章型のデータはどうであろうか。2000年代になるまで、文章型のデータは分析の方法が確立されておらず、多くのデータが蓄積されてきたにも拘わらず、それらのデータは十分な検討がなされないままになっていた。近年、コンピュータによるデータ分析手法が発達し、自由記述のデータの分析方法としてテキストマイニングの手法が使われるようになり（権藤他 2007<sup>4)</sup>、伏木田他 2012<sup>5)</sup>、目久田他 2013<sup>6)</sup>、向井他 2014<sup>7)</sup>）、質的データといわれるものに対して定量的な分析が可能になった。テキストマイニングとは、心理学辞典（2014）<sup>8)</sup> には、「明確で統一的な定義は存在しないが、『大量の文書（テキスト）に記述されている内容の相関や傾向（パターン）を分析する技術』等と定義される。」と記されている。ここで言うテキストは実験協力者が自由記述として書いたものだけではない。インタビュー調査による音声データを文字化し、テキストマイニングによって分析した研究もある（京（2012）<sup>9)</sup>、石館他（2015）<sup>10)</sup>）。このような流れの中で、星野等（2016）<sup>11)</sup> は「言語資源を教育心理学研究においても活用し、言語情報を通して人間の意味の世界を、信頼性の高い方法で検討するような研究の蓄積が待たれるところである。」と述べ、文章型のデータ分析の必要性を示している。これらのことから、子どもの発話データを録音し、それを文字化したものをテキストマイニングによって分析することは可能であり、重要な作業であるように思われる。これまで十分な分析が行われなかった子どもの発話記録についても、テキストマイニングの手法によって再分析を行うことにより、新たな知見が得られるのではないかと考えられる。

神野（吉村の旧姓）（1987）<sup>12)</sup> は、記憶の実験終了後の3歳から9歳までの子ども182名に対し、別室で別の実験者が実験についてのインタビューを行い、その内容を録音し、録音したものを文字化し検討している。この時代にはコンピュータが発達しておらず、全て手作業で、あらかじめ決めておいた特定語に分類される単語の抽出にとどまっている。そこで本研究では、テキストマイニングによる再分析を行い、新たな発見の可能性を検討したい。

## Ⅱ 方法

### 1. 実験計画

2（実験条件）× 5（年齢）の要因計画が用いられた。実験条件の要因は、刺激提示時の言語化の有無による2条件（条件①：言語化なし，条件②：言語化有り）である。年齢の要因は、年少児，年中児，年長児，小学校1年生，小学校3年生である。

### 2. 実験協力者

3歳8か月から9歳8か月の保育園児，幼稚園児，小学生，182名である。内訳は，保育園・幼稚園の年少児40名，年中児40名，年長児40名，小学校1年生24名，3年生38名であった。平均年齢はそれぞれ4歳3ヶ月，5歳3ヶ月，6歳3ヶ月，7歳2ヶ月，9歳3ヶ月であった。

### 3. 手続き

共通体験づくりとそれについて報告するという流れになっている。

#### 3.1 共通体験づくり

1対1で実験が行われた。実験者は実験協力者と机の角をはさんで座り，名前と年齢を問う。実験協力者に目をつぶらせ，その間に台紙の上に刺激を並べる。刺激として色（赤，青，黄）でも，形（正方形，正三角形，丸）でも分けられる9枚のカードが用いられた。正方形の一辺は6 cm，正三角形の高さは6 cm，丸は直径6 cmであった。全てポスターカラーによって彩色された。また，背景を一定にするため，台紙として白色の画用紙が用いられた。画用紙で刺激を隠し，実験協力者の前に置き，目を開けさせる。画用紙を指さし，「ここを見てね。」と言って，画用紙を取る（図1）。

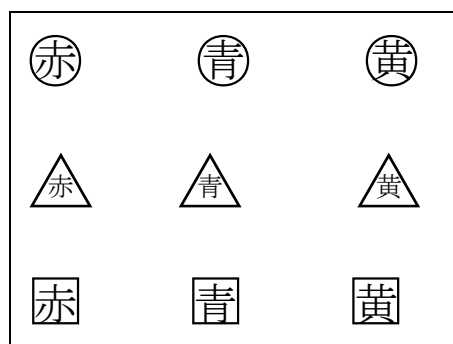


図1 刺激の図形

条件①（言語化無し）では，「これをよく見て，よく覚えてね。」という教示を2度行う。

条件②（言語化有り）では，列および行の類似性について言語化させる。一つの列，または行の3枚のカードを指さしながら，「これと，これと，これはどこが同じかな。」と問う。この問いに答えられない場合，「これと，これと，これは何色（どんな形）かな。」と

問う。この問いにも答えられなかった場合には、色名（または形名）を教える。このやり取りを全ての行、列について行う。言語化課題終了後、「これをよく見て、よく覚えてね」という教示を2度行う。

両条件とも、最長 90 秒間提示する。実験協力者が途中できょろきょろしたり、「覚えた」「わかった」などと言った場合、「もう一度よく見て、よく覚えてね。」と言う。再度、きょろきょろしたり、「覚えた」「わかった」などと言った場合、提示を終了する。

提示が終わると、白画用紙で刺激を隠し、「隠すね」と言い、机の前の方に台紙ごと刺激をずらせ、再生課題を行う。実験協力者の前に白い台紙を置き、実験協力者の左にカードの入った箱を置く。カードの入った箱を指さし、「これを使って、さっき見たのと同じように並べてください。」と言う。実験協力者がカードを並べ終わると、「むずかしかったかな、簡単だったかな。」と問う。

刺激を再度見せ、再生したものと刺激が同じかどうか問う。間違っていると答えた場合には訂正を促す。

### 3.2 報告

記憶の実験終了後、別室で以下の質問に一对一で答える。実験協力者は実験者と机の角をはさんで座る。実験者は会話として不自然にならない程度に無反応であった。

（質問内容）

- ・あそこのお姉さんの所ではどんな事をしたの。
- ・どんな物を使ったの。どんな色だったかな。どんな形だったかな。
- ・あのお姉さんは、〇〇ちゃんにどんな事をしてくださいって言ったの。
- ・できたかな、それともできなかったかな。
- ・面白かったかな、それとも面白くなかったかな。
- ・難しかったかな、それとも簡単だったかな。どうして難しかった（簡単だった）のかな。

## 4. 記録方法

会話は全て小型のカセットテープレコーダーに録音し、これを文字化した。

## 5. 分析方法

テキストマイニングは、文章を自然言語処理してカテゴリ化する部分と、カテゴリ化したデータを統計的に分析する部分に分けることができる。本研究では前者の部分を IBM SPSS Text Analytics for Surveys 4.0 によって行い、後者の部分を IBM SPSS Statistics 22 によって行った。

## III 結果と考察

今回の研究では、「あそこのお姉さんの所ではどんな事をしたの。」という問いに対する反応のみ分析を行った。

## 1. 頻出語

得られた回答文を形態素解析し、よく用いられている語（名詞、動詞、形容詞）を抽出した。分析に使用したソフト（IBM SPSS Text Analytics for Surveys 4.0）では、形態素解析にあたり「キーワード」の抽出を行うが、キーワードには単語だけでなく、複合語も含まれており、「語彙」に近い。表1は名詞、動詞、形容詞ごとに頻出語の一部を示している。各語の頻度は全実験協力者182名中何人がその語を使用したかを示している。

名詞として抽出された語彙数は158であった。「三角」や「丸」、「色」のように刺激材料の特性に当たる単語が多い。「名前」や「目」は実験協力者に名前を問うたこと、目をつぶるように指示したことと関わりがある。

動詞として抽出された語彙数は147であった。「並べる」とい語が多かったが、実験協力者が実験の最後にした具体的な行動なので、印象に残ったと考えられる。「する」については、「並べたりした」のように他の動詞とともに用いられる場合も多いので、反応頻度も多くなった。

形容詞は名詞や動詞と比べて語彙数は少なく26であった。最も頻度が高かったのは「黄色い」であり、刺激の性質を表すものであった。

表1 頻出語

名詞	頻度	動詞	頻度	形容詞	頻度
三角	96	並べる	53	黄色い	16
丸	90	する	32	赤い	4
色	33	覚える	19	四角い	4
同じ	30	ある	18	丸い	4
形	28	それる	17	よい	2
赤	14	やる	13	青い	2
やつ	13	隠す	13	硬い	1
紙	13	置く	12	ない	1
青	12	見る	10	平べったい	1
また	11	遊ぶ	10		
上	7	言う	9		
積み木	6	作る	8		
画用紙	4	わかる	6		
名前	4	違う	4		
目	4	かねる	4		
どこ	4				
あれ	4				
総語彙数	159	総語彙数	147	総語彙数	26

## 2. カテゴリ頻度

テキストマイニングの手法を活用するためには、文章データの中の意味のある語彙に着目し、出現頻度、品詞、類義語、派生語、共起語などの情報をもとにカテゴリを作成する必要がある。そこでまず、文章データに対して形態素解析を行い、キーワード（語彙）を抽出した、次にこれらの語彙の中から意味的に同一の内容を表すと思われる語彙を集め、カテゴリを作成した（表2）。全部で10のカテゴリを作成した。たとえば、「形」のカテゴリには、「三角」、「丸」、「四角い」、「形」などの語彙が含まれている。また、「する」は他の動詞と結合したような場合を除き、「丸した」のように、名詞を直接受ける場合だけを、「する」のカテゴリとした。各カテゴリの頻度は全実験協力者182名中何人がそのカテゴリに言及しているかを示している。

表2 カテゴリ頻度	
カテゴリ	頻度
形	134
色	52
また	11
同じ	36
置く	12
並べる	53
見る	10
隠す	13
覚える	19
する	14

## 3. 各カテゴリと条件との関係

各カテゴリと条件との関連を調べるために $\chi^2$ 検定を行ったが、どのカテゴリについても条件間での違いは見られなかった。

## 4. 各カテゴリと年齢との関係

上述のように条件間で有意な差が見られなかったことから、両条件をまとめて、カテゴリと年齢との関連を $\chi^2$ 検定によって調べた。その結果、「また」、「同じ」、「置く」、「見る」、「隠す」「する」については有意差が見られなかった。これらの言葉は年齢によって出現頻度が変わることはないということが分かった。

そこで、有意差があったもの、有意差はなかったが、意味があると思われる「覚える」について詳述する。

表3は「形」と年齢のクロス表である。 $\chi^2$ 検定の結果、「形」と年齢の関連は有意であった ( $\chi^2=27.102$ ,  $df=4$ ,  $p=.000$ )。年齢が上がるにつれて「形」について言及する子どもが増えることが分かる。

表3 「形」と年齢のクロス表

	年 齢					合計
	年少	年中	年長	小学1年生	小学3年生	
言わない	22	11	9	4	2	48
%	55%	27.5%	22.5%	16.7%	5.3%	26.4%
言った	18	29	31	20	36	134
%	45.0%	72.5%	77.5%	83.3%	94.7%	73.6%
合計	100%	100%	100%	100%	100%	100%

表4は「色」と年齢のクロス表である。 $\chi^2$ 検定の結果、「色」と年齢の関連は有意であった ( $\chi^2=13.805$ ,  $df=4$ ,  $p=.008$ )。年少、年中児では「色」について言及する子どもは少ないが、年長児以上になると「色」について言及するようになることがわかる。また、「形」と「色」を比べると、全体的に「形」について言及する子どもが多い。3歳ではほとんどの子どもが色名を正しく言えるが、この実験事態で「色」について言及をあまりしていないのは、刺激を並べるという行為が、「色」よりも「形」を印象付けたからかもしれない。また、刺激のマトリックス構造(図1参照)を表現できないため、刺激の性質の一面だけに焦点化して報告したようにも思われる。

表4 「色」と年齢のクロス表

	年 齢					合計
	年少	年中	年長	小学1年生	小学3年生	
言わない	35	33	25	13	24	130
%	87.5%	82.5%	62.5%	54.2%	63.2%	71.4%
言った	5	7	15	11	14	52
%	12.5%	17.5%	37.5%	45.8%	36.8%	28.6%
合計	100%	100%	100%	100%	100%	100%

表5は「並べる」と年齢のクロス表である。 $\chi^2$ 検定の結果、「並べる」と年齢の関連は有意であった ( $\chi^2=38.419$ ,  $df=4$ ,  $p=.000$ )。年齢が上がるにつれて「並べる」について言及する子どもが増えることが分かる。この課題の中心的動作は「並べる」ことであり、それを正しく理解し、表現できるようになるのは、小学校1年生以降であることが分かる。

表5 「並べる」と年齢のクロス表

	年 齢					合計
	年少	年中	年長	小学 1 年生	小学 3 年生	
言わない	35	34	33	14	13	129
%	87.5%	85.0%	82.5%	58.3%	34.2%	70.9%
言った	5	6	7	10	25	53
%	12.5%	15.0%	17.5%	41.7%	65.8%	29.1%
合計	100%	100%	100%	100%	100%	100%

表6は「覚える」と年齢のクロス表である。 $\chi^2$ 検定の結果は有意なものではなかったが、クロス表で見ると、小学校 3 年生になるまで、ほとんどの子どもが「覚える」という言葉を言っていないことに意味があるように思われる。実験者は「覚えて下さい」と言っており、子どもたちは「覚える」行為をすることを求められている。子どもたちが「覚える」と言わなかったのは、子どもたちが「覚える」行為をしなかったのか、自分の行った行為は「覚える事」であったと分らないのか、「覚える」という心的動詞を未獲得であるのか、などいくつかの可能性が考えられる。この現象はメタ認知の発達や「心の理論」とも関係していると思われるので更なる研究が必要であると思われる。

表6 「覚える」と年齢のクロス表

	年 齢					合計
	年少	年中	年長	小学 1 年生	小学 3 年生	
言わない	39	40	39	22	23	163
%	97.5%	100.0%	97.5%	91.7%	60.5%	89.6%
言った	1	0	1	2	15	19
%	2.5%	0.0%	2.5%	8.3%	39.5%	10.4%
合計	100%	100%	100%	100%	100%	100%

## 5. 助詞に関する分析

鈴木 (2007) <sup>13)</sup> は、「5 歳半までの幼児は、格助詞『が』と『を』の文法知識において未だ発達過程にある可能性がある。」と述べている。本研究においても格助詞「が」と「を」について検討することにした。IBM SPSS Text Analytics for Surveys 4.0 では、助詞を抽出することはできないので、エクセル上で文字検索を行い「が」と「を」の出現頻度を数えた。「が」は 182 名中 20 名、「を」は 182 名中 48 名であった。

「が」、「を」についても条件間で有意な差が見られなかったことから、両条件をまとめて、「が」、「を」と年齢との関連を  $\chi^2$  検定によって調べた。その結果、「が」については有意差が見られなかった。表 7 は「を」と年齢のクロス表である。 $\chi^2$  検定の結果、「を」と



年齢の関連は有意であった ( $\chi^2=40.247$ ,  $df=4$ ,  $p=.000$ )。年齢が上がるにつれて目的語の前に格助詞「を」付けるようになることが明らかになった。また、半数以上の子どもが「を」を報告場面で使用するようになるのは、小学校3年生であることから、鈴木(2007)で示されている文法理解の年齢よりも、報告のような場面で使用する年齢は高い可能性も考えられる。この点についても検討が必要であろう。

表7 「を」と年齢のクロス表

	年 齢					合計
	年少	年中	年長	小学1年生	小学3年生	
言わない	39	33	29	19	14	134
%	97.5%	82.5%	72.5%	79.2%	36.8%	73.6%
言った	1	7	11	5	24	48
%	2.5%	17.5%	27.5%	20.8%	63.2%	26.4%
合計	100%	100%	100%	100%	100%	100%

#### IV まとめ

本研究の目的は、子どもの発話分析をテキストマイニングの手法のよって行い、言語発達研究方法の可能性を考えることにあった。上記の結果より、子どもの発話データ分析にもテキストマイニングのソフトを用いることが可能であり、多くの知見も得られることが明らかになった。

分析結果から各年齢と発話の特徴を述べる。表8は反応例を示したものであり、表9は無反応または「分からない」と言った年齢別の割合である。

表8 反応例

年齢	反 応 例
年少児	遊んだりした 三角とかね、丸とかね、四角とか
年中児	三角とか四角とか丸とかしてきた
年長児	丸とか三角とか四角で遊んだ 大きな紙に丸とか三角が色で紙の上にのせてあった
小学校1年生	色と形のことしてきた 色板並べてね、勉強ゲームとかした
小学校3年生	形を覚えたり、覚えて並べたりいろんな事を当てたりしました

表9 年齢別無反応または「分からない」と反応した割合

年齢	年少児	年中児	年長児	小学校1年生	小学校3年生
割合	20%	12.5%	7.5%	0%	0%

表9より、小学生になるまでの子どもの中には、無反応、または「分からない」と答えている子どもがいることがわかる。「どんな事をしたの。」という問いに対して、無反応であったり、「わからない」と答えている子どもたちも、「どんな物を使ったの。どんな色だったかな。どんな形だったかな。」と問われると、刺激の特性を答える事が出来た。手続きが複雑で、何が中心的な事であるのか分かりにくく、「何をしたのか」と問われただけでは説明できないようである。

年少児は「形」について言及した子どもは半数以下であり、他の年齢の子どもに比べて大変少ない。また、「形」について言及しても、動詞のない場合が多い(表3, 表8参照)。問われたので、何か答えなくてはならないと感じ、見た物を答えたようである。年中児と年長児はよく似た傾向である。ほとんどの子どもが「形」に言及しており(表3参照)、名詞で終わるのではなく、動詞も付けている(表8参照)。この年齢になると、自分の行動について表現する力がついてくるように思われる。小学校1年生になると、「色」への言及、「並べる」という言葉を使う子どもが増えるが(表4, 表5参照)、格助詞の「を」はあまり使っていない(表7, 表8参照)。発話では目的語と動詞が未分化であるが、書いて答えた場合も同じ傾向であるのか調べてみる必要がある。また、特に分析の対象にはしていないが、「勉強」という言葉を小学1年生の2人だけが言っている。それ以前の子どもでは、「勉強」という言葉が出ていないことから、小学生になり「勉強」と言われることが増え、机に向かうことは、「勉強」と表現されたのかもしれない。小学3年生になると、「並べる」という動詞、格助詞「を」を使用する子どもが増える(表5, 表7参照)。また、心的動容易になるので詞である「覚える」も出てくる(表6, 表8参照)。小学校3年生になると、「何をどうしたのか」という表現ができ、頭の中で行っている事にも言及するようになる。

上述より、子どもの発話もテキストマイニングの手法によって分析でき、発達的变化をとらえることができるということが示された。しかし、この分析で見出されたことは、子どもの言語使用についての発達的变化だけである。次のステップとしては、子どもの言語発達と深い繋がりがあると考えられる他者理解や認知の発達などの側面から、その変化はなぜ起こったのかを検討する必要がある。

#### 引用文献

- 1) 藤崎春代「幼児の報告場面における計画的構成の発達的研究」『教育心理学研究』第30巻第1号, 54 - 63 頁, 1982
- 2) 尾山智子, 仲真紀子「幼児によるポジティブ, ネガティブな出来事の語り: 親が出来事を選定した場合と子どもが出来事を選定した場合」『発達心理学研究』第24巻第1号, 1 -

12 頁, 2013

3) 仲野真史, 長崎勤「幼児におけるナラティブの結束性の発達: ケーキ作り経験に関する報告の分析を通して」『発達心理学研究』第 23 巻第 1 号, 66 - 74 頁, 2012

4) 権藤真織, 呉田陽一「保育学生たちは実習で何を学んだのか? ~実習終了後の総合考察の分析から~」『日本心理学会第 71 回大会論文集』, 2007

5) 伏木田雅子, 北村智, 山内祐平「テキストマイニングによる学部ゼミナールの魅力・不満の検討」『日本教育工学会論文誌』36 巻ショートレター特集, 165 - 168 頁, 2012

6) 目久田純一, 中岡千幸, 越中康治「保育者養成系学科に在学する短期大学生の授業評価基準—テキストマイニングの手法を用いた検討—」『宮城教育大学情報処理センター研究紀要』20 号, 15 - 18 頁, 2013

7) 向井淳治, 徳山絵生, 木本美香, 宮武 望, 小野原未由来, 本荘愛美, 濱田藍子, 高橋直継「テキストマイニングによる病院実務実習実習日誌の分析」『医療薬学』40 巻 4 号, 245 - 251 頁, 2014

8) 家島明彦「社会心理学の質的アプローチ」下山晴彦編集代表『心理学辞典』誠信書房, 285 頁, 2014

9) 京俊輔「保護者と施設職員の『知的障害者の性』に対する意識—テキストマイニングを用いた探索的分析」『島根大学社会福祉論集』4, 1 - 16 頁, 2012

10) 石館美弥子, 山下麻実, いとうたけひこ「小児医療場面において看護師が幼児とのコミュニケーションに用いるオノマトペの特徴」『小児保健研究』第 74 巻第 6 号, 914 - 921 頁, 2015

11) 星野崇宏, 荘島宏二郎, 樋口耕一, 富田英司「教育心理学のためのテキストデータの計量分析」『教育心理学年報』第 55 集, 313 - 321 頁, 2016

12) 神野啓子「課題性の認知に関する発達的研究」『日本心理学会第 51 回大会発表論文集』516 頁, 1984

13) 鈴木孝明「単一項文の理解から探る幼児の格助詞発達」『言語研究』132, 55-76 頁, 2007

## Abstract

The purpose of this study is to investigate language development. A sample of 182 children between the age of 3 and 9 participated in a study of how children talked about experimental situation. Children's accounts were analyzed by text mining method. Our results indicate that there is development in the use of 'shape', 'color', 'put', 'wo(Japanese particle)', and that there are few children, under 8 years old, who use 'remember'.