

作文事例に基づいた児童の「書くこと」に関する学習傾向についての分析

-小学四年生による紹介文・感想文を中心に-

藤田 彬† 田村直良‡

横浜国立大学大学院環境情報学府†

横浜国立大学大学院環境情報研究院‡

E-mail: fujita-akira-rm@ynu.ac.jp† tam@ynu.ac.jp‡

1 はじめに

小学校の国語科の授業における「書くこと」の指導の中では、身の回りで起きた出来事や自己の考えを他者に的確に伝達するための基礎的な技術の学習活動が行われる。この学習活動は主として、児童が作文を記述し、教師による点検（個別の言語要素別の添削と作文に対する総括的な評価）の結果を参照するという枠組を通して行われる。教師による個別の文章表現技術の指導によって、児童は自身の作文に含まれる具体的な問題点を認識し、修正を行うことができる。このような点で、点検は、「書くこと」の能力の育成に重要な働きを持つものと考えられる。可能な限り作文中の全ての問題点に言及し、かつ適切な助言を与える必要があるといえる。しかしながら、表現能力が発達段階にある児童の作文は多様であり、改善を必要とする事項や誤りには、多くのパターンが存在する。これらを俯瞰的に捉えて、全てを把握しておくことは困難であるといえる。これに対して、児童の作文を網羅的に収録したコーパスは、児童の多様な言語活動を把握する上で、重要な役割を果たすと考えられる。

筆者らが知る限り、日本の児童の言語活動をまとめた従来のコーパスは、その多くが話し言葉を収録したコーパスである。一方、永田[1]は児童の書き言葉を収録した「こどもコーパス」を構築し、公開している。小学校5年生3学級81人が書いた、図書をテーマとするブログ形式の文章を1,256編（39,629形態素）収録している。さらに、児童が文章を編集した履歴も同時に収録しており、記述活動の「トレース」が可能という特徴を有している。規模、内容ともに有用性の高いコーパスであるといえる。また、坂本[2]は、国内の小学校のホームページ上に公開された作文を収集し、コーパスを構築した。このコーパスには、265校から収集した10,006編（1,234,961形態素）の作文が収録されている。この収録編数は、それまでに存在した児童の言語活動に関するコーパスと比較して規模が大きい。また、作文の筆者（児童）の学年と性別等の情報も収録されており、有用性の高いものであるといえる。

児童の多様な言語活動を把握することを考えた場合、コーパスには、改善すべき事項の多い作文から優れた作文まで、幅広い能力水準の児童により書かれた作文が収録されることが求められる。また、それぞれの作文の問題点や優れた点に対する指摘事項が収録されることも求められる。

坂本によるコーパスの収集源となった「小学校のWebページに掲載された作文」は、掲載までの過程において何らかの基準で選抜された作文である可能性が高い。少なくとも、指導目標に達しない作文が掲載されることは少ないと考えられる。

また、日本語の児童書き言葉コーパスに関する先行研究には、作文中の個別の箇所に対する添削や、指導事項を基準とする総括的な評価をコーパスに含めたものは、存在しないと考えられる。

本研究では、複数の小学校に在籍する様々な文章能力水準にある児童により、実際のカリキュラムの中で書かれた作文を、収集する。さらに、それらの作文に教師による点検を再現した情報を付与し、コーパスを構築する。また、構築されたコーパスに基づいて児童らの学習状況の分析を行う。なお、本稿において児童とは小学生を指すこととする。

2 児童作文コーパスの構築

2.1 作文事例の収集

神奈川県内9校の小学校から4年生の作文事例を収集した。収集期間は、2011年7月～10月（第Ⅰ期）と12月（第Ⅱ期）の2回に分かれる。第Ⅰ期には9校から620編、第Ⅱ期には第Ⅰ期に収集対象となった小学校のうちの1校から52編、合計672編の作文を収集した。

収集対象とした作文は、本研究のために特別に書かれたものではなく、各校のカリキュラムに沿って、作文活動を行う特定の単元の中で書かれたものである。

収集対象となった全ての小学校が、国語の授業に光村図書出版が発行した教科書を用いている。光村図書出版によ

る4年生の教科書の上編に収録されている物語『一つの花』と、下編に収録されている物語『三つのお願い』を扱う単元の中で書かれた作文を収集した。

『一つの花』は戦時中に生まれた少女とその家族を中心にして戦争体験を描いた作品である。『一つの花』を扱う単元では、感想を交えて作品の内容を他者に紹介する文章（紹介文）が書かれる。

『三つのお願い』は、三つの願いがかなうという一セント硬貨をめぐる、少女と少年の人間関係が描かれた作品である。『三つのお願い』を扱う単元では、作品に対する感想を記述する文章（感想文）が書かれる。

また、教科書の単元に沿った作文とは別に、学年行事である体験学習の後にその感想を記述する感想文も収集した。

表1に、作文の設定と学校別の作文事例の内訳を示す。

作文は、一部を除いて¹、400字詰め原稿用紙に記入されている。これらを複写した後、各校に原本を返却した。個人情報保護の観点から、児童の氏名や本文中の個人を特定する記述（児童の友人の名前、教諭の名前、学校名等）を手でマスキングした上で複写を行った。児童の氏名には筆者の異なりを表現するためのIDを、本文中の個人を特定する記述には任意の名称を、それぞれ代替的に記入した。また、個々の児童が所属する学級についての情報は作文に含まれておらず、学級を区別することはできない。

これらの作文事例をコーパスとして公開するための著作権処理は行っていない。本研究での利用を目的とすることのみについて承諾を得ている。今後、各校と相談した上で公開を検討したい。

2.2 作文の点検

収集した作文事例に対して、15名の人物（以下、点検者）の手で点検を行った。点検者全員が国語教育に関係する人物である。過去に国語教師の経験を持つ者、国語教育系の研究室に所属する大学院生と学部生、塾の講師が含まれる。点検者は、現役の横浜市立小学校教諭1名により、横浜版学習指導要領[3]（特に、小学校第4学年までの「書くこと」の「構成」と「記述」に関する指導事項）に沿った点検方法の研修を事前に受けている。

点検は、点検者が作文を通読した上で、原稿用紙に指摘事項を付記する形で行われた。ここでいう指摘事項とは、児童らが既に指導を受けたと考えられる指導事項と照らし合わせて、不十分あるいは優れていると判断される言語活動に言及する注釈を指す。個別の作文に対する指摘事項

表1 収集した作文事例の内訳（単位：編）

（設定①：『一つの花』の紹介文、設定②：『三つのお願い』の感想文、設定③：体験学習の感想文）

		学校								
		a	b	c	d	e	f	g	h	i
設定	①	131	51	55	84	52	48	63	80	0
	②	0	52	0	0	0	0	0	0	0
	③	0	0	0	0	0	0	0	0	56

の形式には、指摘事項が指し示す範囲について二種の形式がある。

一つは、作文に対する総括的なコメントである。作文全体を通して不十分な点や優れている点、改善につながるアドバイス等を、作文の欄外に付記する。このうち不十分な点と優れている点に関するコメントについては、それぞれ個別の記号を添えることで判別できるようにした。

もう一つは、作文中の個別の箇所に対する注釈である。主に以下に示す事項に該当する箇所について、それを修正する目的のコメントや記号が付記される。

・表記上の誤り

（判読が難しい文字や綴りに誤りのある文字も含む。

ただし、原稿用紙の使い方に関する誤りは含まない。）

例：放課後たども → 放課後なども

例：思ました → 思いました

・句読法上の難点

例：…でした私は… → …でした。私は…

・習得済²の漢字の不使用

例：仲なおり → 仲直り

・文法/語用上の誤り

例：私が思ったのは、…だと私は思いました。

→ 私は…だと思いました。

例：…だということをゆみ子の知っていました。

→ …だということをゆみ子は知っていました。

・主述関係の曖昧性

・段落分けに関する不適切さ

また、誤りを修正するだけでなく、優れている箇所についてのコメントや記号も付記する。

上記の指摘事項とは別に、点検者の総合的な判断による評価を付記した。この総合評価は絶対的な評価でA（良い）、B（標準的）、C（指導の必要がある）の3段階とした。

現時点では、収集した作文事例のうち学校a～fの『一つの花』の紹介文（設定①）421編の点検が完了している。複数人の観点から点検を行う目的で、1編の作文事例につき、異なる2名の点検者がそれぞれ個別に点検を行った。

¹ 学校独自の形式の用紙に書かれた作文が一部ある。この形式の用紙には、マス目がない。

² 全ての作文が小学校4年生により書かれているため、ここでは小学校4年生までに学習する全ての漢字を習得済の漢字とした。

2.3 作文と点検事項の電子コーパス化

作文とそれに対する点検事項をコンピュータによる処理が可能な形に電子化した。基本方針として、収集した児童の作文を可能な限り改変せずにコーパスに収録することを目標とした。点検者による点検事項については、本文中に挿入されるタグ（以下、点検タグ）としてデータ化した。文字コードは UTF-8、改行コードは CR+LF とした。

作文本文の電子化に際して、判読が難しい文字や対応する活字が存在しない文字については、個々の電子化担当者の判断により、推測される文字を充当した。その他の誤りを含む表現や文字列については、改変を加えずにそのまま収録した。また、改行箇所については、行の途中で次行に移るか字下げが見られる箇所を改行箇所として、改行文字を挿入した。段落冒頭への空白文字挿入の有無は、本文中での字下げの有無に従うこととした。

作文の題名も電子化の対象とした。題名は下記のタグを用いて本文と区別する。

<title>...</title>

作文の総合評価についても下記のタグを用いてデータ中に記載する。

<score>...</score>

作文に対する総括的なコメントについての点検タグには、下記の 3 種類がある。

- 優れている点についての講評 <praise>...</praise>
- 不十分な点についての講評 <pointout>...</pointout>
- その他の講評 <comment>...</comment>

本文中の個別の箇所に対する点検タグは、主に注釈が本文に対して及ぼす編集操作の種別に基づいて体系化した。

- 文字列挿入 <ADD seq="", type="">
: 文字列 seq を当該箇所に挿入する。
- 文字列削除 <RM type="">...</RM>
: タグで囲まれた文字列を削除する。
- 文字列変更 <MOD seq="", type="">...</MOD>
: タグで囲まれた文字列を文字列 seq に変更する。
- コメントの付記 <ANT seq="">...</ANT>
: タグで囲まれた文字列にコメント(seq)を付記する。
- 優れた箇所の明示 <GOOD>...</GOOD>
: タグで囲まれた箇所が優れていることを明示する。

<ADD>, <RM>, <MOD>タグは type 属性を持つ。この type 属性は、注釈の内容的な分類を表す属性である。表記上の誤り、句読法上の難点、習得済みの漢字の不使用、文法/語用上の誤りに関する指摘事項の type 属性を他の事項と区別して、"e" という値で示す。現時点では、e 以外の値は設定しておらず、その他の内容の指摘事項については type 属性を null としている。

本文中の個別の箇所に対する点検タグは適用範囲が重なり合うことがあるが、このような場合はタグを入れ子に

表 2 各点検タグの付与事例数の合計

点検タグの種類	付与事例数の合計
<praise>	574
<pointout>	364
<comment>	122
<ADD>	4,222
<RM>	1,876
<MOD>	7,668
<ANT>	2,584
<GOOD>	1,834

して示す。

以下に作文と点検事項のデータ例を示す。

<title>『一つの花』を読んで</title>

<score>B</score>

<pointout>接続詞のあやまりが多い。</pointout>

<praise>感想をたくさん書いている。</praise>

<GOOD><ANT seq="呼びかけの言葉">みなさんは<ADD seq="", type="">『一つの花』という本を<RM type="">、</RM><MOD seq="知", type="e">し</MOD>っていますか。</ANT><GOOD><GOOD><ANT seq="ゆみ子の紹介をしている">主人公はゆみ子という女の子です。</ANT><GOOD> (以下省略)

作文の本文と題名の電子化は全収集事例 672 編について完了している。合計文字数は 322,180 文字、1 編あたりの平均文字数はおよそ 477 文字である。形態素数にして、合計 187,647 形態素、1 編あたりの平均形態素数はおよそ 277 形態素である。形態素数の計数には McCab を用いた。

点検事項の電子化（点検タグの付与）は、全収集事例のうち『一つの花』の紹介文 421 編について完了している。表 2 に点検タグの付与事例数の合計（各作文事例につきそれぞれ 2 名の点検者が付与した点検タグ事例の合計数）をタグの種類別に示す。type 属性が e のタグは全付与事例のうち、4,257 事例存在する。

3 構築したコーパスに基づく児童の学習傾向の分析

前述の通り、構築したコーパスに含まれる点検事項は横浜版学習指導要領に沿ったものである。横浜版学習指導要領の内容は、文部科学省が策定した学習指導要領に沿っているが、指導事項の設定範囲に異なりがある。一般の学習指導要領が指導事項の設定範囲を低学年、中学年、高学年と区別するのに対して、横浜版学習指導要領は指導事項を

表 3 <pointout>タグによる主な指摘内容

タグによって指摘される内容	タグが付与された作文の編数
適切な段落分けができていない	71
自身の考えについての記述が少ない	59
常体と敬体の混用が見られる	30
接続詞の使用法に問題がある	18
句読点の使用法に問題がある	17

1 学年毎に詳細化して設定している。以下では、横浜版学習指導要領に沿った検討を行うこととする。

<pointout>タグは、児童の作文が既習の事柄（カリキュラム上既に指導を受けたと考えられる作文技術）に対して不十分と判断される際に、個別の言語要素に言及した講評を付記するタグである。ここでは、コーパスに付与された<pointout>タグによって指摘される内容を、指導事項に関わる言語要素別に分類して、その付与頻度を分析する。表 3 に、<pointout>タグのうち付与頻度の高い指摘内容を、タグが付与された作文の編数と共に示す。ここでは、2 名の点検者のうちのどちらかがタグを付与した場合に、その作文にタグが付与されたとして、編数を計数した。

「段落分け」と「常体・敬体の使い分け」については、小学校 3 年生で明確に指導事項として設定される内容である。本コーパスの中には、これらの事項を満たしていない作文が高い頻度で見受けられる。

段落分けに関して<pointout>タグが付与された作文は、一切段落分けが行われない作文が大半である。また、段落分けはできていても中心となる段落が明確でない作文が存在する。段落分けには書き手の述べようとすることを明確化する効果があり、「接続詞の使用法」と共に、小学校 5 年生以降で扱われる「事実と意見や感想を分けて書く」技術の基礎となる重要な技法である[4]。

常体・敬体の使い分けの指導は、それぞれの文体が持つ特性を知り、その表現効果を状況別に活かす技術を習得する目的で行われる[4]。文体の使い分けについての通常の指導では、作文が完成した後で児童が混用に気付くことになる。作文の作成中に即時的に混用に対する意識を与えることができれば、より効果的な文体の指導が実現されると考えられる。

「自身の考えを記述すること」について指摘のある作文も多く存在する。これらの作文の大半が、読んだ物語のあらすじや登場人物、時代背景についての客観的な事柄のみを記述したものである。一方で、<praise>タグにより、自身の考えや感想を記述したことに対して優れている旨の講評を付記された作文は 180 編存在する。このことから、自身の考えを記述することを意識する児童とそうでない児童の人数が二極化していると推測される。

「句読点の使用法」については、規範が曖昧であり[5]、適切な使用法を理解できない児童に対して指導を行うことが困難であると考えられる。本コーパスには、<pointout>タグが付与された作文以外にも、句読点の位置や有無を修正する点検事項が多く含まれる。これらの点検事項の持つ情報は、句読法の規範の体系化に役立つものと考えられる。

4 おわりに

本稿では、小学生の作文を収集し、国語教育関係者による点検事項付きのコーパスを構築する手法について述べた。また、構築されたコーパスに基づいた児童らの学習状況の分析結果について述べた。このコーパスに収録される作文は、複数の小学校に在籍する様々な文章能力をもつ児童により、実際のカリキュラムの中で書かれたものである。収集規模も大きく、さらに教師による点検を再現した情報も付与されており、新規性、有用性ともに高い児童作文コーパスであるといえる。

本稿で紹介したコーパスは、筆者らによる作文の自動点検システムの開発に応用されている。今後、収集対象とする学年を広げることで、教育分野に限らず、多くの分野（社会学や発達心理学等）での応用も期待される。また、収集地域の局所性を解消することで、自治体毎の指導目標の異なりに対応したコーパスとする展開も考えられる。さらに、他のドメイン、文種の作文を収集することで、より有用性の高いコーパスとなると考えられる。

今後は、以上のようにコーパスの拡張を進めるとともに、収録作文の著作権処理について検討を行う予定である。

謝辞 本研究については、公益財団法人博報児童教育振興会の児童教育実践事業についての研究助成事業、「学習指導要領に立脚した児童作文自動点検システムの実現」（助成番号：11-B-081、研究代表：藤田彬）の援助を受けている。

参考文献

- [1] 永田亮, 河合綾子, 須田幸次, 掛川淳一, 森広浩一郎. 作文履歴をトレース可能な子供コーパスの構築, 自然言語処理, Vol.17, No.2, pp.51-65(2010).
- [2] 坂本真樹. 小学生の作文コーパスの収集とその応用の可能性. 自然言語処理, Vol.17, No.5, pp.75-98(2010).
- [3] 横浜市教育委員会. 横浜版学習指導要領 国語科編, 株式会社ぎょうせい(2009).
- [4] 国語教育研究所. 作文技術指導大辞典, 明治図書(1996).
- [5] 比毛博. 点のうち方, ことばの科学 5, pp.221-237(1992).